(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 August 2001 (02.08.2001)

PCT

(10) International Publication Number
WO 01/55442 A2

(51) International Patent Classification[6]: C12P 21/06

(21) International Application Number: PCT/US01/02543

(22) International Filing Date: 25 January 2001 (25.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/491,404	25 January 2000 (25.01.2000)	US
09/617,746	17 July 2000 (17.07.2000)	US
09/631,451	3 August 2000 (03.08.2000)	US
09/663,870	15 September 2000 (15.09.2000)	US
60/250,583	30 November 2000 (30.11.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	09/491,404 (CIP)
Filed on	25 January 2000 (25.01.2000)
US	09/617,746 (CIP)
Filed on	17 July 2000 (17.07.2000)
US	09/631,451 (CIP)
Filed on	3 August 2000 (03.08.2000)
US	09/663,870 (CIP)
Filed on	15 September 2000 (15.09.2000)
US	60/250,583 (CIP)
Filed on	30 November 2000 (30.11.2000)

(71) Applicant *(for US only)*: **HYSEQ, INC.** [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **FORD, John, E.** [US/US]; 2763 South Norfolk, #210, San Mateo, CA 94403 (US). **BOYLE, Bryan, J.** [US/US]; 1947 10th Avenue, San Francisco, CA 94116 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **LIU, Chenghua** [CN/US]; 1125 Ranchero Way, Apt. #14, San Jose, CA 95117 (US). **ASUNDI, Vinod** [US/US]; 709 Foster City Boulevard, Foster City, CA 94404 (US). **ZHOU, Ping** [CN/US]; 7595 Newcastle Drive, Cupertino, CA 95014 (US). **ZHANG, Jie** [CN/US]; 4930 Poplar Terrace, Campbell, CA 95008 (US). **DRMANAC, Radoje, T.** [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with declaration under Article 17(2)(a); without abstract; title not checked by the International Searching Authority*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/55442 A2

(54) Title: NOVEL BONE MARROW NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract:

NOVEL BONE MARROW NUCLEIC ACIDS AND POLYPEPTIDES

1. BACKGROUND OF THE INVENTION

1.1 TECHNICAL FIELD

The present invention provides novel bone marrow-expressed polynucleotides and bone marrow-expressed proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

1.2 BACKGROUND

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

The bone marrow is a well-organized tissue located within the central cavity of bone. It has a complex three-dimensional structure that is richly innervated and highly vascularized. Two primary cell types make up the structure of the bone marrow. These are the stromal, and parenchymal cells. Stromal cells include reticular cells such as fibroblasts, endothelial cells, adipocytes, as well as cells of the osteochondrogenic lineage. They exert important influences on osteoclastogenesis and lymphopoiesis, and have additional effects on bone turnover. Parenchymal cells are comprised of the hematopoietic cells, and are important for proliferation, maturation, and migration of cells that make up the blood.

In the adult, hematopoiesis takes place primarily in the bone marrow. Therefore, all of the cells that make up the blood, such as erythrocytes, platelets, basophils, natural killer cells, eosinophils, T- and B-lymphocytes, neutrophils, macrophages, and others, are produced in this structure. Each of these cells is derived from a common, self-renewing stem cell that

proliferates, and/or differentiates depending on regulatory molecules that are produced by the stromal cells. Stromal cells are predominantly a mixture of fibroblasts, macrophage/dendritic lineage cells, epithelial cells, and endothelial cells. They influence the fate of hematopoietic cells through the secretion of soluble factors, cytokines, and the expression of membrane-anchored growth factors, and cell surface recognition molecules.

Cytokines are necessary for normal hematopoiesis in the bone marrow, and provide a means of fine-tuning bone marrow function in response to stimulation. They are not only produced by stromal cells, but can also be secreted by macrophages, and antigen-stimulated T lymphocytes for the purpose of replenishing leukocytes that may be consumed during immune and inflammatory reactions. Many cytokines that influence the differentiation and expansion of hematopoietic progenitor cells are termed colony-stimulating factors, because they were initially assayed by their ability to stimulate the formation of cell colonies in bone marrow cultures. Some of these colony-stimulating factors (CSFs) include, granulocyte-CSF, granulocyte/macrophage-CSF, monocyte-CSF, Kit-ligand, interleukin (IL)-6, FLK-2 ligand, and leukemia inhibitory factor. Each of these stimulates the growth and development of various leukocytic or erythroid colonies. Other cytokines secreted in the bone marrow include IL-9, a T cell line and mast cell progenitor-stimulating factor, IL-11, a megakaryocytopoiesis stimulator, and IL-7, a cytokine that influences the survival and expansion of immature precursors committed to the B and T cell lineages. Many other cytokines are also secreted in the bone marrow.

Cell-surface molecules that represent several adhesion molecule superfamilies including integrins, selectins, sialomucins and the immunoglobulin domain-containing proteins, are important in supporting cell-cell and cell-extracellular matrix interactions in the bone marrow. These proteins are critical to the homing of progenitor cells selectively to the marrow stroma for proliferation and differentiation. They also serve to influence the fate of the progenitor cells by directing them to differentiate into a specific lineage. For example, VLA-4 directs control of late erythroid differentiation and pro-B cell maturation.

The bone marrow is also the site of B cell development. B cells begin as lymphoid stem cells that differentiate into progenitor B-cells, or pro-B cells. Pro-B cells proliferate within the bone marrow, and fill the extravascular spaces between large sinusoids in the shaft of the bone. They next mature into precursor B cells, pre-B cells. The stromal cells of the bone marrow are crucial for both pro- and pre-B cell development because they provide a source of cytokines, and a substrate for direct interaction with the pro- and pre-B cells. Pro-B cells require interaction with VCAM-1 and stem-cell factor (SCF) on the stromal cells to induce expression of the IL-7 receptor. Secretion of IL-7 by the stromal cells then induces the pro-B cells to mature into pre-B

cells. Continued IL-7 secretion by stromal cells induces pre-B cells to begin proliferating and eventually differentiates them into immature B-cells. In addition, a selection process within the bone marrow eliminates B cells with self-reactive phenotypes, functioning to protect against autoimmune disease.

The bone marrow environment also influences bone turnover and bone precursor cell functions. Bone marrow stromal cells include the precursors of the osteochondrogenic lineage, and can modulate the effects of some systemic factors on bone turnover. Furthermore, hematopoietic cells may influence the differentiation of osteogenic cells, and mature lymphocytes may impact osteoclastic and osteoblastic functions. For instance, B-lymphocytes have been implicated in the secretion of factors that change the immunological milieu at sites of new bone induction and influence new bone formation.

The identified bone marrow-expressed polynucleotide and polypeptide sequences may have applications in hematopoiesis, stem cell survival, and bone growth and remodeling. Identification of secreted factors that stimulate hematopoiesis may serve to produce greater immune responses in immunosuppressed individuals. The identification of factors that preferentially stimulate specific hematopoietic cell types may also allow the prevention of specific disorders such as anemia in the case erythroid cell stimulating factors, or platelet deficiency in the case of megakaryocyte stimulating factors. Likewise, stem cell stimulating factors may be used to restore blood cell populations following chemotherapy treatments for cancer. Therapy to stimulate bone healing and remodeling may also be identified by the discovery of novel factors in the bone marrow that influence bone resorption by osteoclasts, or new bone cell differentiation from stromal cells.

2. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides from bone marrow tissue, and novel isolated polynucleotides from bone marrow tissue encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by

hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-87, 175-261 or 349-353 and are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenine; C is cytosine; G is guanosine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO: 1-87, 175-261 or 349-353 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO: 1-87, 175-261 or 349-353. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO: 1-87, 175-261 or 349-353 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO: 1-87, 175-261 or 349-353. The sequence information can be a segment of any one of SEQ ID NO: 1-87, 175-261 or 349-353 that uniquely identifies or represents the sequence information of SEQ ID NO: 1-87, 175-261 or 349-353.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information are provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-87, 175-261 or 349-353, or novel segments or parts of the nucleic acids of the invention are used as primers in

expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-87, 175-261 or 349-353 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying bone marrow tissues and cells; for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO: 1-87, 175-261 or 349-353; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO: 1-87, 175-261 or 349-353; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-87, 175-261 or 349-353. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO: 1-87, 175-261 or 349-353; (b) a nucleotide sequence encoding any one of the amino acid sequences comprising 88 – 174, 262-348 or 354-358 as set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO: 1-87, 175-261 or 349-353; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions, or (c) polypeptides comprising any of the polypeptide sequences set forth in SEQ ID NO: 88-174, 262-348 or 354-358. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention. The polypeptides may have the initial methionine (Met) removed.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, *e.g., in situ* hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (*e.g.*, bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provide methods for treatment that involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

affect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Tables 1A-D and 7); for which they have a signature region (as set forth in Table 2 and 8); or for which they have homology to a gene family (as set forth in Table 3). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in increasing hematopoiesis, stem cell survival, and bone growth and remodeling.

3. DETAILED DESCRIPTION OF THE INVENTION

3.1 DEFINITIONS

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide that retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line

stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived. The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides that modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs is nucleic acid fragments that induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonucleotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30

nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs: 1-87, 175-261 or 349-353.

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NOs: 1-87, 175-261 or 349-353. The sequence information can be a segment of any one of SEQ ID NOs: 1-87, 175-261 or 349-353 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO: 1-87, 175-261 or 349-353. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4^{20} possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosome. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match ($1 \div 4^{25}$) times the increased probability for mismatch at each nucleotide position (3 x 25). The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence that encodes for the full-length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence that encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell that removes any leader/signal sequence. The mature protein portion may or may not include an initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant"(or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, *e g.*, recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes that produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 Daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (*e.g.*, microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, *e.g., E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells that have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells that have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (*e.g.*, soluble proteins) or partially (*e.g.*, receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

As used herein, "substantially equivalent" or "substantially similar" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (*i.e.*, the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, *e.g.*, mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, *e.g.*, mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% sequence identity, more preferably at least 98% sequence identity, and most preferably at least 99% sequence identity. Substantially equivalent nucleotide sequence of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, the nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least 85% sequence identity, more preferably at least 90%, sequence identity, more preferably at least 95% sequence identity, more preferably at least 98% sequence identity, and most preferably at least 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (*e.g.*, via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J. (1990) Methods Enzymol. 183:626-645). Identity between

sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides that mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

3.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO: 1-87, 175-261 or 349-353; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO: 88-174, 262-348 or 354-358; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polynucleotides of any one of SEQ ID NO: 1-87, 175-261 or 349-353. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO: 1-87, 175-261 or 349-353; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 88-174, 262-348 or 354-358. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in

receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO: 1-87, 175-261 or 349-353 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO: 1-87, 175-261 or 349-353 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO: 1-87, 175-261 or 349-353 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, *e.g.*, at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99% sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO: 1-87, 175-261 or 349-353 or complements thereof, which fragment is greater

than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that are selective for (i.e. specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided in SEQ ID NO: 1-87, 175-261 or 349-353, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NOs: 1-87, 175-261 or 349-353 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NOs: 1-87, 175-261 or 349-353, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST, which stands for Basic Local Alignment Search Tool, is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm could also be used.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences that encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid

sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, *e.g.*, by substituting first with conservative choices (*e.g.*, hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (*e.g.*, hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent

amino acid sequence may be used in the practice of the invention for the cloning and expression of these novel nucleic acids. Such DNA sequences include those that are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO: 1-87, 175-261 or 349-353, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NOs: 1-87, 175-261 or 349-353 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NOs: 1-87, 175-261 or 349-353 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise

regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBS, phagescript, PsiX174, pBluescript SK, pBS KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination

signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

3.3 ANTISENSE NUCLEIC ACIDS

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1-87, 175-261 or 349-353, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of

SEQ ID NO: 1-87, 175-261 or 349-353 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO: 1-87, 175-261 or 349-353 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences that flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding a nucleic acid disclosed herein (*e.g.*, SEQ ID NO: 1-87, 175-261 or 349-353, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of an mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of an mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of an mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine; uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -uni ts, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

3.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of an mRNA. A ribozyme having specificity for a nucleic acid of the invention can

be designed based upon the nucleotide sequence of a DNA disclosed herein (*i.e.*, SEQ ID NO: 1-87, 175-261 or 349-353). For example, a derivative of Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (*e.g.*, promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, Proc. *Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

3.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of bone marrow-derived DNA sequences allows for modification of cells to permit, or increase, expression of bone marrow-derived polypeptide. Cells can be modified (e.g.,

by homologous recombination) to provide increased bone marrow-derived polypeptide expression by replacing, in whole or in part, the naturally occurring bone marrow specific promoter with all or part of a heterologous promoter so that the cells express bone marrow-derived polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to bone marrow-derived polunucleotide-encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the bone marrow-derived polynucleotide-coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the bone marrow-derived polynucleotide- coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural levels. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a

compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5′ flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability

of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

3.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO: 88 – 174, 262-348 or 354-358 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NOs: 1-87, 175-261 or 349-353 or the corresponding full length or mature protein. Polypeptides

of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NOs: 1-87, 175-261 or 349-353 or (b) polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO: 88 – 174, 262-348 or 354-358 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO: 88 – 174, 262-348 or 354-358 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO: 88 – 174, 262-348 or 354-358.

Fragments of the proteins of the present invention that are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which it is expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments that differ from a

nucleic acid fragment of the present invention (*e.g.*, an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells that have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell that produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells that naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography,

and immuno-affinity chromatography. See, *e.g.*, Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: *A Laboratory Manual*; Ausubel et al., *Current Protocols in Molecular Biology*. Polypeptide fragments that retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays that are well known in the art to identify molecules that bind to the polypeptides. These molecules include but are not limited to, for e.g., small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO: 88 – 174, 262-348 or 354-358.

The protein of the invention may also be expressed as a product of transgenic animals, *e.g.*, as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, *e.g.*, U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the

importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, Calif., U.S.A. (the MaxBat™ kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl™ or Cibacrom blue 3GA Sepharose™; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form, which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLabs (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP- HPLC) steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially

homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, e.g., antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties, which may be fused to the polypeptide, include therapeutic agents that are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

3.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Mol. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMATRIX software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), PFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference), SignalP software package (Nielsen H et al., Int. J. Neural Syst., Vol. 8, pp. 581 – 599 (1997), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol. Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for

Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

3.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus, or to the middle.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprise one or more domains fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to

avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention

3.8 GENE THERAPY

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be

inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell, which drives expression of the polynucleotides in the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences that affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences that alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both

upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques that can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

3.9 TRANSGENIC ANIMALS

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

3.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

3.10.1 RESEARCH USES AND UTILITIES

The research community can use the polynucleotides provided by the present invention for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch

and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

3.10.2 NUTRITIONAL USES

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

3.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation,

Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

3.10.4 STEM CELL GROWTH FACTOR ACTIVITY

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells *in vivo* or *ex vivo* is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce

large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, bone marrow inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be

used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support e.g. as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

3.10.5 HEMATOPOIESIS REGULATING ACTIVITY

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with

irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/bone marrows (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

3.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention that induces cartilage and/or bone growth in circumstances where bone is not normally formed has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of

tendonitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further, conditions that may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

3.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders that may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (*e.g.*, anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66,

1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process that requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven

Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self-tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of auto-reactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of auto-reactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to

reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in

Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad. Sci. USA 88:7548-7551, 1991.

3.10.8 ACTIVIN/INHIBIN ACTIVITY

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A

polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

3.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146,

1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

3.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke)).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

3.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell

cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine. Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g.

exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These *in vitro* models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

3.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1- 7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988;

Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

3.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science 282*:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol*, 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol*, 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem*, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

3.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number

of different libraries used for the identification of compounds, and in particular small molecules, that modulate (*i.e.*, increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) is then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications i.e. phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

3.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflammation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

3.10.16 LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

3.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

(i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions that sever a portion of the nervous system, or compression injuries;

(ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

(iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human

immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;

(iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;

(v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;

(vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;

(vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and

(viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

(ii) increased sprouting of neurons in culture or in vivo;

(iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or

(iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, *etc.*,

depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, *e.g.*, weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

3.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

3.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA, which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

3.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis are determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et al., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a

suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

3.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

3.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01μg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1μg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

3.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents that either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti- inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, *e.g.*, treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

3.12.1 ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral

ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

3.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations that can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the

pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic,

talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations that can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances that increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents that increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium

carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B-lymphocytes will respond to antigen through their surface immunoglobulin receptor. T-lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments that the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention

and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μg to about 100 mg (preferably about 0.1 μg to about 10 mg, more preferably about 0.1 μg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention that are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention that may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above-mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns.

In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly (ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly (vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, *e.g.*, amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also affect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a

mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

3.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD_{50} and ED_{50}. Compounds that exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, *e.g.*, Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety that are sufficient to maintain the

desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen that maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μg/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μg/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

3.12.4 PACKAGING

The compositions may, if desired, be presented in a pack or dispenser device that may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

3.13 ANTIBODIES

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen-binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab}, $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well,

such as IgG_1, IgG_2, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, such as an amino acid sequence shown in SEQ ID NO: 88 – 174, 262-348 or 354-358, and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of TGF alpha-like protein that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

The term "specific for" indicates that the variable regions of the antibodies of the invention recognize and bind polypeptides of the invention exclusively (*i.e.*, able to distinguish the polypeptide of the invention from other similar polypeptides despite sequence identity, homology, or similarity found in the family of polypeptides), but may also interact with other

proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow et al. (Eds), Antibodies A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the polypeptides of the invention are also contemplated, provided that the antibodies are first and foremost specific for, as defined above, full-length polypeptides of the invention. As with antibodies that are specific for full length polypeptides of the invention, antibodies of the invention that recognize fragments are those which can distinguish polypeptides from the same family of polypeptides despite inherent sequence identity, homology, or similarity found in the family of proteins.

Antibodies of the invention are useful for, for example, therapeutic purposes (by modulating activity of a polypeptide of the invention), diagnostic purposes to detect or quantitate a polypeptide of the invention, as well as purification of a polypeptide of the invention. Kits comprising an antibody of the invention for any of the purposes described herein are also comprehended. In general, a kit of the invention also includes a control antigen for which the antibody is immunospecific. The invention further provides a hybridoma that produces an antibody according to the invention. Antibodies of the invention are useful for detection and/or purification of the polypeptides of the invention.

Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

The labeled antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays to identify cells or tissues in which a fragment of the polypeptide of interest is expressed. The antibodies may also be used directly in therapies or other diagnostics. The present invention further provides the above-described antibodies immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and Sepharose®, acrylic resins and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir, D.M. et al., "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific

Publications, Oxford, England, Chapter 10 (1986); Jacoby, W.D. et al., Meth. Enzym. 34 Academic Press, N.Y. (1974)). The immobilized antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays as well as for immuno-affinity purification of the proteins of the present invention.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

3.13.1 POLYCLONAL ANTIBODIES

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface-active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants that can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

3.13.2 MONOCLONAL ANTIBODIES

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen-binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal

Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

3.13.3 HUMANIZED ANTIBODIES

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539). In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues that are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

3.13.4 HUMAN ANTIBODIES

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by

transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al, (Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals that are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse™ as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells that secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least

one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

3.13.5 FAB FRAGMENTS AND SINGLE CHAIN ANTIBODIES

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

3.13.6 BISPECIFIC ANTIBODIES

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two

immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers that are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific

antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., J. Immunol. 147:60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc RI (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a

radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

3.13.7 HETEROCONJUGATE ANTIBODIES

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

3.13.8 EFFECTOR FUNCTION ENGINEERING

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

3.13.9 IMMUNOCONJUGATES

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from

Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

3.14 COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium that can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the

presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (*e.g.* text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NOs: 1-87, 175-261 or 349-353, or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NOs: 1-87, 175-261 or 349-353 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein-encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present

invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs that are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence that match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software include, but are not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids; more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration that is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

3.15 TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA. Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan

et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

3.16 DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers,

Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container, which will accept the test sample, a container, which contains the antibodies used in the assay, containers, which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers, which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats that are well known in the art.

3.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection).

See, e.g., Kunkel et al., U.S. Pat. No. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide *in vivo* at the target site.

3.18 SCREENING ASSAYS

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide set forth in SEQ ID NO: 88 – 174, 262-348 or 354-358 encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NOs: 1-87, 175-261 or 349-353, or which binds to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

(a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and

(b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds that modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds that modulate the expression of a polynucleotide of the

invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs that rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives that have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems.

Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents that bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents that bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

3.19 USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NOs: 1-87, 175-261 or 349-353. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NOs: 1-87, 175-261 or 349-353 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample. Preferably a hybridization probe from any of nucleotide sequences SEQ ID NO: 1-87, 175-261 or 349-353 can be used as an indicator of bone marrow tissue.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well-known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

3.20 PREPARATION OF SEQUENCING CHIPS AND ARRAYS

A basic example is using 6-mers attached to 50 micron surfaces to give a chip with dimensions of 3 x 3 mm which can be combined to give an array of 20 x 20 cm. Another example is using 9-mer oligonucleotides attached to 10 x 10 microns surface to create a 9-mer chip, with dimensions of 5 x 5 mm. 4000 units of such chips may be used to create a 30 x 30 array. In an array in which 4,000 to 16,000 oligochips are arranged into a square array. A plate, or collection of tubes, as also depicted, may be packaged with the array as part of the sequencing kit.

The arrays may be separated physically from each other or by hydrophobic surfaces. One possible way to utilize the hydrophobic strip separation is to use technology such as the Iso-Grid Microbiology System produced by QA Laboratories, Toronto, Canada.

Hydrophobic grid membrane filters (HGMF) have been in use in analytical food microbiology for about a decade where they exhibit unique attractions of extended numerical range and automated counting of colonies. One commercially available grid is ISO-GRID™ from QA Laboratories Ltd. (Toronto, Canada) which consists of a square (60 x 60 cm) of polysulfone polymer (Gelman Tuffryn HT-450, .45 um pore size) on which is printed a black hydrophobic ink grid consisting of 1600 (40 x 40) square cells. HGMF have previously been inoculated with bacterial suspensions by vacuum filtration and incubated on the differential or selective media of choice.

Because the microbial growth is confined to grid cells of known position and size on the membrane, the HGMF functions more like an MPN apparatus than a conventional plate or membrane filter. Peterkin et al. (1987) reported that these HGMFs can be used to propagate and store genomic libraries when used with a HGMF replicator. One such instrument replicates growth from each of the 1600 cells of the ISO-GRID and enables many copies of the master HGMF to be made (Peterkin et al., 1987).

Sharpe et al. (1989) also used ISO-GRID HGMF form QA Laboratories and an automated HGMF counter (MI-100 Interpreter) and RP-100 Replicator. They reported a technique for maintaining and screening many microbial cultures.

Peterkin and colleagues later described a method for screening DNA probes using the hydrophobic grid-membrane filter (Peterkin et al., 1989). These authors reported methods for effective colony hybridization directly on HGMFs. Previously, poor results had been obtained due to the low DNA binding capacity of the epoxysulfone polymer on which the HGMFs are printed. However, Peterkin et al. (1989) reported that the binding of the DNA to the surface of the membrane was improved by treating the replicated and incubated HGMF with polyethyleneimine, a polycation, prior to contact with DNA. Although this early work uses cellular DNA attachment, and has a different objective to the present invention, the methodology described may be readily adapted for Format 3 SBH.

In order to identify useful sequences rapidly, Peterkin et al. (1989) used radiolabeled plasmid DNA from various clones and tested its specificity against the DNA on the prepared HGMFs. In this way, DNA from recombinant plasmids was rapidly screened by colony hybridization against 100 organisms on HGMF replicates that can be easily and reproducibly prepared.

Manipulation with small (2-3 mm) chips, and parallel execution of thousands of the reactions. The solution of the invention is to keep the chips and the probes in the corresponding arrays. In one example, chips containing 250,000 9-mers are synthesized on a silicon wafer in the form of 8 x 8 mM plates (15 uM/oligonucleotide, Pease et al., 1994) arrayed in 8 x 12 format (96 chips) with a 1 mM groove in between. Probes are added either by multichannel pipette or pin array, one probe on one chip. To score all 4000 6-mers, 42 chip arrays have to be used, either using different ones, or by reusing one set of chip arrays several times.

In the above case, using the earlier nomenclature of the application, F=9; P=6; and F+P=15. Chips may have probes of formula BxNn, where x is a number of specified bases B; and n is a number of non-specified bases, so that x= 4 to 10 and n= 1 to 4. To achieve more efficient hybridization, and to avoid potential influence of any support oligonucleotides, the specified bases can be surrounded by unspecified bases, thus represented by a formula such as (N)nBx(N)m.

3.21 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, e.g., Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen *et al.*, (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via a phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm$_7$), is then added to a final concentration of 10 mM 1-MeIm$_7$. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm$_7$, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) Proc. Natl. Acad. Sci. USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

3.22 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA,

including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *Cvi*JI, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease *Cvi*JI normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (*Cvi*JI**), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a *Cvi*JI** digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that *Cvi*JI** restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

3.23 PREPARATION OF DNA ARRAYS

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm^2, depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane. Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm^2 and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments that are intended as illustrations of single aspects of the invention, and compositions and methods that are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of

the present preferred embodiments. Consequently, the only limitations that should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

4.0 EXAMPLES

4.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosomes using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences that flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones from each cluster were selected for sequencing.

The sequence of the amplified inserts, in some cases, was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems (ABI) sequencer to obtain the novel nucleic acid sequences.

4.2 EXAMPLE 2

Novel Nucleic Acids

The novel nucleic acids of the present invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The nucleic acids of SEQ ID NO: 1-87, inclusive, were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend some of the seed ESTs into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 119, gb pri 119, and UniGene version 119, Geneseq October version, and Genscan, Genemark and Hyseq gene predictions on human genomic sequence from the human genome project updated October 2000) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

4.3 EXAMPLE 3

Further Characterization

Clusters from Example 1 were identified which were expressed in bone marrow tissue cDNA libraries, but not in other tissues. Novel nucleic acids were assembled by the method of Example 2. A subset of the assembled nucleic acids comprising sequences from the identified clusters was selected. This subset includes SEQ ID NO: 1-87. The tissue sources in which SEQ ID NO: 1-87 were exclusively expressed were found to be in BMD001 and BMD002 bone marrow libraries (Clontech).

The homologies for SEQ ID NO: 1-87 and the corresponding peptide sequences, SEQ ID NO: 88 – 174 were obtained by performing various searches as shown in Tables 1A to 1D and as discussed herein.

The homologous sequences to SEQ ID NO: 88-174 were obtained by a BLASTP version 2.0al 19MP-WashU search against the Geneseq database updated November 9, 2000, update 23 for year 2000 (Derwent), using the BLAST algorithm. The homologues for SEQ ID NO: 88-174 from Geneseq are shown in Table 1A below.

The homologous sequences to SEQ ID NO: 88-174 were also obtained by a BLASTP version 2.0al 19MP-WashU search against the NCBI Genbank nr database updated November 10, 2000, using the BLAST algorithm. The homologues for SEQ ID NO: 88-174 from Genbank are shown in Table 1B below.

The homologous sequences to SEQ ID NO: 1-87 were also obtained by a BLASTN version 2.0al 19MP-WashU search against the Geneseq database updated November 9, 2000, update 23 for year 2000 (Derwent), using the BLAST algorithm. The homologues for SEQ ID NO: 1-87 from Geneseq are shown in Table 1C below.

The homologous sequences to SEQ ID NO: 1-87 were also obtained by a BLASTN version 2.0al 19MP-WashU search against the NCBI Genbank nt database updated November 10, 2000, using the BLAST algorithm. The homologues for SEQ ID NO: 1-87 from Genbank are shown in Table 1D below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), the polypeptide sequences corresponding to SEQ ID NO: 1-87 and 349-353 were examined to determine whether they had identifiable signature regions. Table 2 and 8 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the PFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) the polypeptide sequences corresponding to SEQ ID NO: 1-87 and 349-353 were examined for domains with homology to certain peptide domains. Table 3 shows the name of the domain found, the description, the e-value and the PFam score for the identified domain within the sequence.

The polypeptide sequence within each of SEQ ID NO: 88-174 that is the predicted signal peptide sequence and its cleavage site can be determined using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A mean S score, as described in the Nielson et. al. was obtained for the polypeptide sequences. Table 4 shows the position of the predicted signal peptide in each of the polypeptides and the mean score associated with that signal peptide.

Table 5 provides a correlation between the amino acid sequences set forth in the sequence listing, the nucleotide sequence encoding the amino acid sequence, the corresponding contig nucleotide sequence and amino acid sequence determined by the method of Example 4 and the corresponding full length edited sequence determined by the method of Examples 5 and 6.

4.4 EXAMPLE 4

Assemblage of Novel Nucleic Acids

The contigs or nucleic acids of the present invention, designated as SEQ ID NO: 175-261 were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

The nucleotide sequence within the assembled contigs that codes for signal peptide sequences and their cleavage sites was determined from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the

publication "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, vol. 10, no. 1, pp.1-6 (1997) incorporated herein by reference,. A maximum S score and a mean S score, as described in the Nielson et al. reference, are obtained from each assembled contig. Table 6 sets forth the nucleotide range for each sequence of SEQ ID NO: 262-348 that encodes a corresponding amino acid sequence containing the signal peptide sequence and its cleavage site, the maximum S score and the mean S score obtained for each sequence.

A signal peptide or leader peptide is usually a segment of about 15 to 30 amino acids at the N terminus of protein that enables the protein to be targeted to a cell membrane or secreted from a cell. Generally, the signal peptide acts as an export label and is removed as the protein is secreted in its final form.

4.5 EXAMPLE 5

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e., dbEST version 117, gb pri 117, UniGene version 117, Genpept release 117). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide sequences, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS: 349 - 351. The amino acids are SEQ ID NO: 354-356.

The homology for SEQ ID NO: 349 - 351 were obtained by a BLASTP version 2.0a1 19MP-WashU search against Genpept release 118, using BLAST algorithm. The results showed homologues for SEQ ID NO: 349-351 from Genpept. The homologues with identifiable functions for SEQ ID NO: 349-351 are shown in Table 7 below.

4.6 EXAMPLE 6

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e., dbEST version 118, gb pri 118, UniGene version 118, Genpept release 118). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-

ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide sequence, including splice variants resulting from these procedures is shown in the Sequence Listing as SEQ ID NO: 352 and 353. The amino acid is SEQ ID NO: 357 and 358.

The homology for SEQ ID NO: 352 and 353 was obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 118, using BLAST algorithm. The results showed homologues for SEQ ID NO: 352 and 353 from Genpept. The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologues with identifiable functions for SEQ ID NO: 352 and 353 are shown in Table 7 below.

The nucleotide sequence within the sequences the codes for signal peptide sequences and its cleavage site can be determined using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A mean S score, as described in the Nielson et. al. was obtained for the polypeptide sequences. Table 9 shows the position of the predicted signal peptide in each of the polypeptides and the mean score associated with that signal peptide.

TABLE 1A

SEQ ID NO:	ACCESSION NO.	BLAST SCORE	P-VALUE	% IDENTITY	DESCRIPTION
6	Y41768	2360(835.8bits)	5.7e-245	93	Y41768 Human PRO286 protein sequence. Length = 1041
11	Y05069	1039(370.8bits)	5.5e-105	97	Y05069 Human PIGR-2 protein sequence. Length = 205
14	Y73475	135(52.6bits)	3.4e-09	52	Y73475 Human secreted protein clone ye7_1 protein sequence SEQ ID NO:172.
22	W40481	256(95.2bits)	2.4e-21	39	W40481 Human SH2 binding protein. Length = 503
74	R28150	138(53.6bits)	1.2e-08	41	R28150 Sugar beet chitinase 1. Length = 439

TABLE 1B

SEQ ID NO:	ACCESSION NO.	BLAST SCORE	P-VALUE	% IDENTITY	DESCRIPTION
5	NP_055659.1	389(142.0bits)	1.9e-34	87	KIAA0329 gene product [Homo sapiens] >dbj|BAA20787.1| (AB002327) KIAA0329 [Homo sapiens]
6	AAF78036.1	2364(837.2bits)	3.6e-245	93	AF245703_1 (AF245703) toll-like receptor 8 [Homo sapiens] Length = 1041
7	AAF89682.1	537(194.1bits)	1.5e-51	98	(AF169968) DNA binding protein DESRT [Mus musculus] Length = 743
11	NP_006669.1	410(149.4bits)	4.2e-38	47	CMRF35 leukocyte immunoglobulin-like receptor; CMRF35 antigen [Homo sapiens] >sp|Q08708|CM35_HUMAN CMRF35 ANTIGEN
13	BAB15450.1	649(233.5bits)	2.0e-63	96	(AK026332) unnamed protein product [Homo sapiens] Length = 258
22	CAB96828.1	618(222.6bits)	3.8e-60	99	(AL109658) dJ776F14.2 (novel immunoglobulin domains containing protein) [Homo sapiens]
74	CAB62280.1	148(57.2bits)	1.5e-09	41	(AJ242540) hydroxyproline-rich glycoprotein DZ-HRGP [Volvox carteri f. nagariensis]

TABLE 1C

SEQ ID NO:	ACCESSION NO.	BLAST SCORE	P-VALUE	% IDEN-TITY	DESCRIPTION
1	Z41384	874(137.2bits)	2.3e-34	65	Z41384 Human normal uterus tissue derived cDNA 60. Length = 1346
2	C32570	1175(182.3bits)	3.0e-47	94	C32570 Human secreted protein 5' EST, SEQ ID NO: 36645. Length = 263
3	Z52488	5037(761.8bits)	3.1e-302	96	Z52488 Human secreted protein clone ye2_1 nucleotide sequence SEQ ID NO:27.
6	Z34304	14243(2143.1bits)	0.0e+00	97	Z34304 Human PRO286 nucleotide sequence. Length = 4199
7	C09016	1460(225.1bits)	3.8e-59	81	C09016 Human secreted protein 5' EST, SEQ ID NO: 13091. Length = 470
8	Z24819	1502(231.4bits)	1.5e-62	99	Z24819 Human secreted protein gene 9 clone HMCFY13. Length = 883
10	A16628	2828(430.4bits)	5.1e-123	98	A16628 Human secreted protein clone ci52_2 nucleotide sequence SEQ ID NO:21.
11	X28250	2885(438.9bits)	4.0e-125	94	X28250 Human PIGR-2 coding sequence. Length = 1137
12	X83003	683(108.5bits)	1.9e-24	63	X83003 Human WRN genomic sequence. Length = 87,350
14	Z52560	1914(293.2bits)	1.5e-81	80	Z52560 Human secreted protein clone ye7_1 nucleotide sequence SEQ ID NO:171.
16	Z42365	1079(167.9bits)	6.8e-43	95	Z42365 Human 5' EST isolated from a cDNA library SEQ ID NO:124. Length = 249
18	X40516	1547(238.2bits)	3.7e-64	96	X40516 Human secreted protein 5' EST SEQ ID No: 116. Length = 334
21	T35033	362(60.4bits)	3.3e-09	72	T35033 HSV-1 IR-L (position 117156-117341) target sequence. Length = 186
22	A45057	333(56.0bits)	1.2e-06	57	A45057 Mouse secreted expressed sequence tag SEQ ID NO:1632. Length = 676
33	Z97376	328(55.3bits)	2.9e-08	73	Z97376 Human prostate cancer differentially expressed gene #237. Length = 483
37	X83003	349(58.4bits)	2.4e-09	63	X83003 Human WRN

SEQ ID NO:	ACCESSION NO.	BLAST SCORE	P-VALUE	% IDEN-TITY	DESCRIPTION
					genomic sequence. Length = 87,350
45	V88906	647(103.1bits)	8.7e-24	66	V88906 EST clone IA188. Length = 634
50	Z00475	506(82.0bits)	1.3e-16	71	Z00475 Human secreted protein cDNA endoding gene 66. Length = 2483
52	V88316	1365(210.9bits)	2.7e-56	99	V88316 EST clone EZ676. Length = 756
53	A42278	817(128.6bits)	5.7e-31	92	A42278 Human secreted expressed sequence tag SEQ ID NO:1018. Length = 205
54	X84971	585(93.8bits)	7.2e-21	73	X84971 Human secreted protein gene No. 39. Length = 1176
62	V86232	578(92.8bits)	2.0e-20	71	V86232 EST clone S70. Length = 362
63	C19606	1559(240.0bits)	1.1e-64	97	C19606 Human secreted protein 5' EST, SEQ ID NO: 23681. Length = 324
65	X79021	290(49.6bits)	4.1e-07	68	X79021 Human secreted protein gene 11 clone HSLCU73. Length = 798
74	Q76213	459(74.9bits)	6.5e-12	60	Q76213 HSV L/ST region. Length = 12,001
77	Z10752	560(90.1bits)	7.0e-19	63	Z10752 Genomic sequence of the human HKNG1 gene. Length = 72,604
85	C20378	480(78.1bits)	4.3e-15	82	C20378 Human secreted protein 5' EST, SEQ ID NO: 24453. Length = 176
87	Z14832	986(154.0bits)	9.0e-39	99	Z14832 Human gene expression product cDNA sequence SEQ ID NO:2301. Length = 300

SEQ ID NO:	ACCESSION NO	BLAST SCORE	P-VALUE	% IDENTITY	DESCRIPTION
1	AL138836.15	4093(620.2bits)	1.9e-177	96	AL138836 Human DNA sequence from clone RP11-88M19 on chromosome 9, complete sequence [Homo sapiens]
2	AK026357.1	3756(569.6bits)	6.5e-247	93	AK026357 Homo sapiens cDNA: FLJ22704 fis, clone HSI12602 Length = 1984
3	AK000399.1	10017(1509.0bits)	0.0	98	AK000399 Homo sapiens cDNA FLJ20392 fis, clone KAIA4653 Length = 2091
4	AK000812.1	732(115.9bits)	2.6e-26	98	AK000812 Homo sapiens cDNA FLJ20805 fis, clone ADSE02009 Length = 1586
5	AL137229.2	6231(941.0bits)	6.2e-318	99	CNS01DWK Human chromosome 14 DNA sequence *** IN PROGRESS *** BAC C-2246N19 of library CalTech-D from chromosome 14 of Homo
6	AF246971.1	15253(2294.6bits)	0.0	97	AF246971 Homo sapiens Toll-like receptor 8 (TLR8) mRNA, complete cds
7	AF169968.1	5798(876.0bits)	4.9e-256	82	AF169968 Mus musculus DNA binding protein DESRT (Desrt) mRNA, complete cds
10	AL023285.1	812(127.9bits)	3.1e-29	73	HS474A14 Human DNA sequence from clone 474A14 on chromosome 1q24.1-25.2 Contains EST, CA repeat, 5'UTR (tenascin-R), GSS,
11	X66171.1	611(97.7bits)	3.5e-32	63	HSCMRF35A H.sapiens CMRF35 mRNA, complete CDS Length = 1151
12	AL049745.9	4714(713.3bits)	5.3e-210	96	HSJ654H19 Human DNA sequence from clone 654H19 on chromosome 1p31.1-33 Contains ESTs, STSs, GSSs and CpG Islands, complete
13	AK026332.1	3269(496.5bits)	2.3e-282	98	AK026332 Homo sapiens cDNA: FLJ22679 fis, clone HSI10687 Length = 2587
14	AL137521.1	6823(1029.8bits)	0.0	95	HSM802253 Homo sapiens mRNA; cDNA DKFZp434D0218 (from clone DKFZp434D0218); partial cds
16	AC007565.1	2717(413.7bits)	9.3e-183	96	AC007565 Homo sapiens chromosome 19, cosmid R27656, complete sequence
18	AC002553.1	5064(765.9bits)	1.4e-277	94	AC002553 Homo sapiens chromosome 17, clone hCIT529I10, complete sequence
21	M24972.1	379(62.9bits)	5.1e-09	72	DDICTSRE D.discoideum CT-rich satellite rDNA, clone pCT8 Length = 210
22	AL109658.5	3342(507.5bits)	7.3e-233	94	HSJ776F14 Human DNA sequence from clone RP4-776F14 on chromosome

SEQ ID NO:	ACCESSION NO	BLAST SCORE	P-VALUE	% IDENTITY	DESCRIPTION
					20p12.2-13 Contains the 5' end of the FKBP1A gene for
23	AL109658.5	1768(271.3bits)	2.0e-72	97	HSJ776F14 Human DNA sequence from clone RP4-776F14 on chromosome 20p12.2-13 Contains the 5' end of the FKBP1A gene for
26	AC006370.2	4056(614.6bits)	9.1e-176	94	AC006370 Homo sapiens BAC clone RP11-292P9 from Y, complete sequence
31	AC005803.1	1790(274.6bits)	2.1e-73	100	AC005803 Homo sapiens chromosome 17, clone hRPK.214_C_8, complete sequence
33	AL035209.1	3943(597.7bits)	1.1e-170	98	HS7H11 Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 7H11, complete sequence [Homo sapiens]
37	AC004605.1	434(71.2bits)	3.7e-12	66	HUAC004605 Homo sapiens Chromosome 16 BAC clone CIT987SK-A-248F7, complete sequence
38	AL109753.9	3954(599.3bits)	3.7e-171	98	HSJ875J14 Human DNA sequence from clone RP5-875J14 on chromosome Xq13.1-21.1, complete sequence [Homo sapiens]
41	AC009116.7	1452(223.9bits)	3.8e-58	91	AC009116 Homo sapiens chromosome 16 clone RP11-477D3, complete sequence
43	AC004973.1	357(59.6bits)	7.0e-06	60	AC004973 Homo sapiens PAC clone RP5-1139I1 from Xq23, complete sequence
45	AL031287.3	3877(587.8bits)	1.1e-167	98	HS703H14 Human DNA sequence from clone 703H14 on chromosome 1q23.2-24.3 Contains 3' end of a novel gene, ESTs, CA
47	AC006038.2	3139(477.0bits)	2.4e-134	98	AC006038 Homo sapiens BAC clone RP11-299C5 from 2, complete sequence
49	AL121760.11	3123(474.6bits)	1.2e-133	97	HSDJ968J1 Human DNA sequence from clone RP5-968J1 on chromosome 20 Contains part of a novel gene similar to collagen
50	AK002135.1	506(82.0bits)	1.2e-15	71	AK002135 Homo sapiens cDNA FLJ11273 fis, clone PLACE1009338 Length = 2067
54	AC004782.1	1131(175.7bits)	1.2e-43	73	AC004782 Homo sapiens chromosome 5, BAC clone 205e20 (LBNL H170), complete sequence
55	AF125358.1	1940(297.1bits)	9.2e-86	95	F125350S09 Homo sapiens cytohesin 1 gene, exon 10, alternatively spliced
60	AL158141.14	1753(269.1bits)	9.8e-72	92	AL158141 Human DNA sequence from clone RP11-351K23 on chromosome X,

SEQ ID NO:	ACCESSION NO	BLAST SCORE	P-VALUE	% IDENTITY	DESCRIPTION
					complete sequence [Homo sapiens]
62	X03205.1	578(92.8bits)	5.4e-19	71	HSRRN18S Human 18S ribosomal RNA Length = 1869
64	Z98052.1	2326(355.0bits)	1.3e-97	98	HS505B13 Human DNA sequence from clone 505B13 on chromosome 1p36.2-36.3 Contains CA repeat and GSSs, complete sequence [Homo
65	AC005630.1	1353(209.1bits)	1.1e-53	86	AC005630 Homo sapiens PAC clone RP5-1129D5 from 15, complete sequence
68	AC023510.16	1921(294.3bits)	2.5e-79	99	AC023510 Homo sapiens 12 BAC RP11-713N11 (Roswell Park Cancer Institute Human BAC Library) complete sequence
69	AC006430.22	1290(199.6bits)	8.0e-51	97	AC006430 Homo sapiens chromosome 9, clone RP11-525G7, complete sequence
70	AC004764.1	586(94.0bits)	5.0e-19	66	AC004764 Homo sapiens chromosome 5, P1 clone 255g5 (LBNL H61), complete sequence
72	AL078605.30	1626(250.0bits)	5.3e-66	91	HSJ894D12 Human DNA sequence from clone RP5-894D12 on chromosome 6q26-27. Contains part of the gene for a novel protein
74	AJ250235.1	2945(447.9bits)	1.2e-125	85	HSA250235 Homo sapiens FECH gene for ferrochelatase, exons 1-11 Length = 38,637
77	AC004470.1	589(94.4bits)	3.7e-19	66	AC004470 Homo sapiens Xp22 BAC GSHB-433024 (Genome Systems Human BAC library) complete sequence
78	AL121945.6	1687(259.2bits)	5.0e-69	99	HSDJ352G1 Human DNA sequence from clone RP3-352G1 on chromosome 6q21-22.2 Contains a GSS, complete sequence [Homo sapiens]
79	AL049715.25	610(97.6bits)	4.2e-20	74	HSJ646P11 Human DNA sequence from clone RP4-646P11 on chromosome 1, complete sequence [Homo sapiens]
80	AC005226.1	1573(242.1bits)	1.3e-63	98	AC005226 Homo sapiens PAC clone RP4-683L10 from 14q24.3, complete sequence
84	AL354829.8	1608(247.3bits)	3.4e-65	92	AL354829 Human DNA sequence from clone RP11-218B22 on chromosome 13, complete sequence [Homo sapiens]
85	AC021068.17	2772(422.0bits)	9.1e-118	96	AC021068 Homo sapiens 3 BAC RP11-480O22 (Roswell Park Cancer Institute Human BAC Library) complete sequence
87	AC009484.3	348(58.3bits)	2.8e-08	60	AC009484 Homo sapiens BAC

SEQ ID NO:	ACCESSION NO	BLAST SCORE	P-VALUE	% IDENTITY	DESCRIPTION
					clone RP11-313O11 from 2, complete sequence

Table 2

SEQ ID NO:	ACCESSION NO:	DESCRIPTION	P-VALUE	RAW SCORE	RESIDUE POSITION
6	PR00019B	LEUCINE-RICH REPEAT SIGNATURE	9.100e-10	11.36	26-40
6	PR00019B	LEUCINE-RICH REPEAT SIGNATURE	1.000e-09	11.36	102-116
6	PR00019B	LEUCINE-RICH REPEAT SIGNATURE	8.560e-09	11.36	239-253
8	PD01171A	COAT PROTEIN GLYCOPROTEIN VP2 V.	6.459e-09	14.02	12-38
11	DM01688I	2 POLY-IG RECEPTOR.	7.480e-10	14.97	70-118
11	DM01688J	2 POLY-IG RECEPTOR.	4.455e-09	14.69	23-60
13	BL00227E	Tubulin subunits alpha, beta, and gamma proteins.	9.695e-10	24.15	137-172
15	BL01288C	Uncharacterized protein family UPF0027 proteins.	7.500e-09	10.54	34-41
22	DM00031B	IMMUNOGLOBULIN V REGION.	5.215e-09	15.41	70-104
22	PD02870D	RECEPTOR INTERLEUKIN-1 PRECURSOR.	8.755e-09	15.74	86-121

Table 3

SEQ ID NO	ACCESSION NO.	PFAM MODEL NAME	PFAM SCORE	E-VALUE
6	LRR	PF00560	19.2	0.096
6	LRR	PF00560	9.0	48
6	LRR	PF00560	17.9	0.25
6	LRR	PF00560	18.0	0.22
6	LRR	PF00560	13.1	6.6
6	LRR	PF00560	10.7	27
6	LRR	PF00560	6.2	1.2e+02
6	LRR	PF00560	12.0	15
6	LRR	PF00560	5.5	1.6e+02
6	LRR	PF00560	0.1	9.7e+02
6	LRR	PF00560	8.1	66
11	ig	PF01812	27.1	7.4e-07
22	ig	PF01812	38.6	2.1e-10

TABLE 4

SEQ ID NO:	SIGNAL PEPTIDE POSITION	MEAN SCORE	CUTOFF	CONCLUSION
1	1-45	0.104	0.48	NO
2	1-63	0.176	0.48	NO
3	1-177	0.194	0.48	NO
4	1-29	0.604	0.48	YES
5	1-483	0.404	0.48	NO
6	1-123	0.105	0.48	NO
7	1-498	0.061	0.48	NO
8	1-43	0.120	0.48	NO
9	1-15	0.057	0.48	NO
10	1-53	0.330	0.48	NO
11	1-192	0.211	0.48	NO
12	1-25	0.444	0.48	NO
13	1-141	0.370	0.48	NO
14	1-83	0.346	0.48	NO
15	1-16	0.724	0.48	YES
16	1-75	0.242	0.48	NO
17	1-11	0.173	0.48	NO
18	1-37	0.534	0.48	YES
19	1-12	0.704	0.48	YES
20	1-22	0.618	0.48	YES
21	1-18	0.714	0.48	YES
22	1-43	0.076	0.48	NO
23	1-63	0.229	0.48	NO
24	1-26	0.816	0.48	YES
25	1-69	0.104	0.48	NO
26	1-17	0.743	0.48	YES
27	1-22	0.856	0.48	YES
28	1-18	0.777	0.48	YES
29	1-330	0.105	0.48	NO
30	1-22	0.897	0.48	YES
31	1-34	0.202	0.48	NO
32	1-73	0.227	0.48	NO
33	1-40	0.346	0.48	NO
34	1-53	0.215	0.48	NO
35	1-28	0.747	0.48	YES
36	1-26	0.246	0.48	NO
37	1-18	0.226	0.48	NO
38	1-25	0.850	0.48	YES
39	1-22	0.612	0.48	YES
40	1-44	0.152	0.48	NO
41	1-20	0.667	0.48	YES
42	1-22	0.902	0.48	YES
43	1-47	0.075	0.48	NO
44	1-26	0.147	0.48	NO
45	1-18	0.885	0.48	YES
46	1-40	0.410	0.48	NO
47	1-16	0.969	0.48	YES
48	1-14	0.775	0.48	YES
49	1-25	0.626	0.48	YES
50	1-33	0.571	0.48	YES
51	1-71	0.122	0.48	NO
52	1-33	0.758	0.48	YES
53	1-22	0.943	0.48	YES
54	1-37	0.618	0.48	YES
55	1-53	0.085	0.48	NO
56	1-19	0.939	0.48	YES
57	1-20	0.815	0.48	YES

58	1-16	0.866	0.48	YES
59	1-29	0.380	0.48	NO
60	1-16	0.141	0.48	NO
61	1-36	0.581	0.48	YES
62	1-56	0.345	0.48	NO
63	1-43	0.664	0.48	YES
64	1-16	0.416	0.48	NO
65	1-18	0.769	0.48	YES
66	1-62	0.514	0.48	YES
67	1-25	0.490	0.48	YES
68	1-29	0.556	0.48	YES
69	1-38	0.656	0.48	YES
70	1-40	0.821	0.48	YES
71	1-22	0.782	0.48	YES
72	1-28	0.411	0.48	NO
73	1-38	0.778	0.48	YES
74	1-62	0.152	0.48	NO
75	1-45	0.780	0.48	YES
76	1-23	0.652	0.48	YES
77	1-15	0.818	0.48	YES
78	1-17	0.764	0.48	YES
79	1-40	0.204	0.48	NO
80	1-18	0.667	0.48	YES
81	1-23	0.845	0.48	YES
82	1-34	0.318	0.48	NO
83	1-0	0.000	0.48	NO
84	1-43	0.419	0.48	NO
85	1-40	0.107	0.48	NO
86	1-21	0.693	0.48	YES
87	1-88	0.208	0.48	NO

TABLE 5

SEQ ID NO of nucleo-tide sequence	SEQ ID NO: of peptide sequence	SEQ ID NO: of contig nucleo-tide sequence	SEQ ID NO: of contig peptide sequence	SEQ ID NO: of full-length nucleo-tide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO in USSN 09/491,404 of contig nucleotide sequence	Priority Docket NO: Corresponding SEQ ID NO: in priority application	SEQ ID NO. of nucleotide sequence in USSN 60/250,583	SEQ ID NO. of peptide sequence in USSN 60/250,583
1	88	175	262			279		2574	11
2	89	176	263			1709		2587	24
3	90	177	264			1602		2603	40
4	91	178	265			3095		2621	58
5	92	179	266			453		2654	91
6	93	180	267			1149		2672	109
7	94	181	268			2976		2784	221
8	95	182	269	351	356	3792	785CIP2B_214	2826	263
9	96	183	270			2385		2834	271
10	97	184	271	353	358	2871	785CIP2C_23	2837	273
11	98	185	272			3565		2846	282
12	99	186	273			1081		2853	289
13	100	187	274	352	357	610	785CIP2C_2	2859	295
14	101	188	275			2155		2866	302
15	102	189	276			329		2871	307
16	103	190	277			1480		2872	308
17	104	191	278			2262		2877	313
18	105	192	279			823		2890	326
19	106	193	280			1681		2902	338
20	107	194	281			1310		2906	342
21	108	195	282			2478		2907	343
22	109	196	283			3612		2925	361
23	110	197	284			3612		2926	362
24	111	198	285			168		2947	383
25	112	199	286			1083		2949	385
26	113	200	287			550		2950	386
27	114	201	288			1082		2951	387
28	115	202	289			919		2952	388
29	116	203	290			826		2957	393
30	117	204	291			1372		2972	408
31	118	205	292			1808		2974	410
32	119	206	293			252		2984	420
33	120	207	294			1373		2987	423
34	121	208	295			3572		2990	426
35	122	209	296			2229		2995	431
36	123	210	297			2235		3003	439
37	124	211	298			1527		3004	440
38	125	212	299			3293		3006	442
39	126	213	300			2232		3007	443
40	127	214	301			780		3017	453
41	128	215	302			2234		3060	496
42	129	216	303			875		3061	497
43	130	217	304			2409		3066	502
44	131	218	305	350	355	3097	785CIP2B	3069	505

SEQ ID NO of nucleo-tide sequence	SEQ ID NO: of peptide sequence	SEQ ID NO: of contig nucleo-tide sequence	SEQ ID NO: of contig peptide sequence	SEQ ID NO: of full-length nucleo-tide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO in USSN 09/491,404 of contig nucleotide sequence	Priority Docket NO: Corresponding SEQ ID NO: in priority application	SEQ ID NO. of nucleotide sequence in USSN 60/250,583	SEQ ID NO. of peptide sequence in USSN 60/250,583
							_192		
45	132	219	306			917		3081	517
46	133	220	307			1309		3104	540
47	134	221	308			2304		3109	545
48	135	222	309	349	354	2391	785CIP2B _64	3113	549
49	136	223	310			1051		3123	559
50	137	224	311			3538		3148	584
51	138	225	312			3289		3155	591
52	139	226	313			2149		3156	592
53	140	227	314			3301		3163	599
54	141	228	315			176		3193	629
55	142	229	316			2393		3209	645
56	143	230	317			3093		3224	660
57	144	231	318			2228		3316	752
58	145	232	319			482		3357	793
59	146	233	320			2387		3360	796
60	147	234	321			3184		3402	838
61	148	235	322			3532		3412	848
62	149	236	323			3538		3429	865
63	150	237	324			2054		3628	1064
64	151	238	325			3758		3632	1068
65	152	239	326			1700		3675	1111
66	153	240	327			3292		3702	1138
67	154	241	328			2123		3706	1142
68	155	242	329			2219		3761	1197
69	156	243	330			3209		3768	1204
70	157	244	331			1123		3790	1226
71	158	245	332			3383		3838	1274
72	159	246	333			3530		3923	1359
73	160	247	334			874		3948	1384
74	161	248	335			2152		4143	1578
75	162	249	336			2390		4212	1647
76	163	250	337			3385		4257	1692
77	164	251	338			175		4271	1706
78	165	252	339			868		4380	1814
79	166	253	340			2221		4387	1821
80	167	254	341			1360		4453	1886
81	168	255	342			3384		4530	1961
82	169	256	343			2227		4555	1985
83	170	257	344			315		4591	2021
84	171	258	345			1766		4722	2151
85	172	259	346			3103		4847	2275
86	173	260	347			1682		5012	2440
87	174	261	348			2153		5126	2554

Table 6

SEQ ID NO:	Predicted beginning nucleotide location corresponding to first amino acid residue of amino acid sequence	Predicted end nucleotide location correspondi ng to first amino acid residue of amino acid sequence	Amino acid segment containing signal peptide (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
262	33	239	MCIELCIIFPVVMSSPSFFILVSENNNNDLFLIL MVLLLVSLGLTLAVLVDWLQLEDLKMASPIC LGP*
263	2971	3261	MMQLLFPLPLWGIIPFHLHCCDIVCPLSQVEG GILRLPPALVHSIFLLHAACVMISCKAFNIKSP LCVRPHVERKTCLREEVCCVSPFSSPQICVS*
264	1095	1244	MSRFFIFCCLRHFSYFSDFAVLFLGALEHLKY QLAVGHSVLSESTDGNV*
265	540	388	MAYRGQLLAGFTFDVSACLWTSWRTALTEC VAWGICPLGWVVPVLGPVDG*
266	129	320	MHRGVLVTLLKITVLKSMHRGILVTLLKITIL KSMHRGVLDTLLKITILKSMHRGVLVTLLKIT
267	3	608	MVMLAALAHHLFYWDVWFIYNVCLAKVKG YRSLSTSQTFYDAYISYDTKDASVTDWVINEL RYHLEESRDKNVLLCLEERDWDPGLAIIDNL MQSINQSKKTVFVLTKKYAKSWNFKTAFYLA LQRLMDENMDVIIFILLEPVLQHSQYLRLRQR ICKSSILQWPDNPKAEGLFWQTLRNVVLTEN DSRYNNMYVDSIKQY*
268	954	694	MLFWLIKVSCSFSCSDETSAASWGFGAFSFSF LLLGISCLMRLVPDTFVLFSFSCELFSCFRGLI GGRGLSSSPLINLSYGRINLS*
269	107	1135	MLQGHSSVFQALLGTFFTWGMTAAGAALVF VFSSGQRRILDGSLGFAAGVMLAASYWSLLA PAVEMATSSGGFGAFAFFPVAVGFTLGAAFV YLADLLMPHLGAAEDPQTALALNFGSTLMK KKSDPEGPALLFPESELSIRIGRAGLLSDKSEN GEAYQRKKAAATGLPEGPAVPVPSRGNLAQP GGSSWRRIALLILAITIHNVPEGLAVGVGFGAI EKTASATFESARNLAIGIGIQNFPEGLAVSLPL RGAGFSTWRAFWYGQLSGMVEPLAGVFGAF AVVLAEPILPYALAFAAGAMVYVVMDDIIPE AQISGNGKLASWASILGFVVMMSLDVGLG*
270	160	357	MKCKLIPVCPFLRLNTQPLLIISYGIFLHIFRDF SYIHRVRERHSVFLSVGQQWCPELTRSIFLLN
271	147	773	MGLGARGAWAALLLGTLQVLALLGAAHESA AMAASANIENSGLPHNSSANSTETLQHVPSD HTNETSNSTVKPPTSVASDSSNTTVTTMKPTA ASNTTTPGMVSTNMTSTTLKSTPKTTSVSQNT SQISTSTMTVTHNSSVTSAASSVTITTTMHSEA KKGSKFDTGSFVGGIVLTLGVLSILYIGCKMY YSRRGIRYRTIDEHDAII*

SEQ ID NO:	Predicted beginning nucleotide location corresponding to first amino acid residue of amino acid sequence	Predicted end nucleotide location correspondi ng to first amino acid residue of amino acid sequence	Amino acid segment containing signal peptide (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
272	120	434	MWLPPALLLLSLSGCFSIQGPESVRAPEQGSL TVQCHYKQGWETYIKWWCRGVRWDTCKILI ETRGSEQGEKSDRVSIKDNQKDRTFTVTMEG LRRDDAXVYWCG
273	14	196	MKFHLSFFSLKRAIFYICAKADKISGGYLYKC RTVSYSGKNVRSGVKISGFLSACIISYL*
274	140	586	MHMLNGALLALLFPVVNTRLLPFELEIYYIQH VMLYVVPIYLLWKGGAYTPEPLSSFRWALLS TGLMFFYHFSVLQILGLVTEVNLNNMLCPAIS DPFYGPWYRIWASGHQTLMTMTHGKLVILFS YMAGPLCKYLLDLLRLPAKKID*
275	1095	322	MRWIAFAVMIVLALIRIGHGQGEGHPPLADFS GVRNLFGVCVYSFMCQHSLPSLITPVSSKRHL TRLVFLDYVLILAFYGLLSFTAIFCFRGDSLM DMYTLNFARCDVVGLAAARLFLGLFPVFTIST NFPIIAVTLRNNWKTLFHREGGTYPWVVDRV VFPTITLVPPVLVAFCTHDLESLVGITGAYAG TGIQYVIPAFLVYHCRRDTQLAFGCGVSNKH RSPFRHTFWVGFVLLWAFSCFIFVTANIILSET KL*
276	35	337	MALALAAYVCGWVVDRETWPVPMPCNKGG RACNLEMGMEWLNLHCEVSKWQQPPSGALC CSLAPLQSIFFPAAKVMFKNGSWTVLLPCSEF PIGFPSHLE*
277	706	951	MRCGWGPLGCLGTGAPAGWMVLGSPRSQLQ RARWSRASLSAFGWEIRLRPEGPKAPRQLLL VALESETLGVHGGATPLHCL*
278	79	273	MNNSPLALFSWEGWKKFLVLLPAFCITPSQST SFSNIVPTTYQYCTPGSCQAVHSNAVGGNTW K*
279	1121	1387	MFSCFFSTSLATSVSLEAQSCFAWPLIVSFPQG SLLSPFLLMSYNLSHLIYSGELNGRLYAENSQI CICSPAFPTKLYLHIFADLITS*
280	12	182	MCLAHLFKLLVYFNRSNSWVQAPFVLETTTG LFSSSVSLICILNLFCKQNLNNNFL*
281	125	319	MLPPLCWCCVRTMTCCIGTSTGMDGRPPSPW RRIPCWTQTCSCRNSATPSSPHFLHTSRWPGP MY
282	6	239	MTARFLICLFQTTMYAEFNLGQRRWQTRNAP NLSGWLGLAGAAPWQGRISPMLGTKVSLCN LSEESLAPLAKHTPRA*
283	695	877	MSPTGLLVVFAPVVLGLKAITLAALLLALATS RRSPGQEDVKTTGPAGAMNTLAWSKGQE*

SEQ ID NO:	Predicted beginning nucleotide location corresponding to first amino acid residue of amino acid sequence	Predicted end nucleotide location correspondi ng to first amino acid residue of amino acid sequence	Amino acid segment containing signal peptide (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
284	695	877	MSPTGLLVVFAPVVLGLKAITLAALLLALATS RRSPGQEDVKTTGPAGAMNTLAWSKGQE*
285	241	438	MFFWRLLILYSSPEITVCLHLFTSSGLKMQHL RSHCHLFRQALFLCFSDTTVMGFFLSYWWQF SV*
286	130	279	MWSVTSTIFICRLLIVRLLGNTAVRTSVVFLPH KAGRHWEKSTSLVSGG*
287	88	276	MWADSILASLLLWPHQSLQLWHHPHLANKN MGVPPPTTCKPWSTVAQKFADYIPFMTTWPP LG
288	25	534	MRLLHCKTLHIVLFTLLYKILMDHQNLSEHV LCMVLYLIELGLENSAEEESDEEASVGGPERC HDSWFPGSNLVSNMRHFINYVRVRVPETAPE VKRDSPASTSSDNLGSLQNSGTAQVFSLVAER RKKFQEIINRSSSEANQVVRPTTSSKWSAPGS APQLTTAIFGN*
289	102	308	MKMFQMLLTSSFCSLSHLQSCQHISFLSISNHS KIFLYLQPTCYLYLPPLPLFSRSWHWNLRVHI CSP*
290	785	1090	MCVAACFSLVAWSILQWGKRKYPEGNSSWQ IKEKVWRFSTAFCSVNEWKFADILSMADHLK KCSYNVVEKREEAIPLPCMCVTRELTKEGRSL RSVLKPVL*
291	737	940	MCTFRGLLTGLLTFPLFSPVLYFCNKFPNKTN MFLLCFCKNYFLSTVFFIFLRQSFVLVAQTGV QGV*
292	342	656	MKGILFFFFWKGVYFSPSLKPRGEIWVNCPQP WGEGGPIGGKIKNGGVFSGREFFPTMEKKKF PPRAKTKINPPRKMGAQRRPTPKWPTRQGPF NRSPKKGKRYP
293	79	363	MPWVLGCTPFIALAYFFLWFLPPFTSLRGLW YTTFYCLFQALATVPYTALTMLLTPCPRERDS ATAYRMTVEMAGTLMGATVHGLIVSGAHRP HR
294	279	434	MAVEPLLAHFLRWSWLSARDFYSLGNVDPA LWVPCFFLLFLLIITDNNNDS*
295	196	351	MCSVTCGVLFALSGLLLYSSPSPHWNRPSRIA VYLMCLTKYCTGSSAASCQ*
296	460	630	MPACCYRPCLLQPISLLNILLLLMRKPSQEVIN DTPKAGKWLSRYLDSGLFYSCACG
297	119	382	MWSWHVQLQVSAPLHHLLCLHFPPAHRIYM PFPSPKRAPAMLNKGIHMQGMSSVSWKGEA KFSFHHQRVAFNIIYTRQAFALLVLLN*

SEQ ID NO:	Predicted beginning nucleotide location corresponding to first amino acid residue of amino acid sequence	Predicted end nucleotide location correspondi ng to first amino acid residue of amino acid sequence	Amino acid segment containing signal peptide (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
298	88	243	MQVSGPRPQLFLPSVFFVLLFSYTFTETTQWT VVILALNSKLSFKEIETIF*
299	78	263	MFMPGTVLRILLALPYLILTKQVQFFLFSDEIM AWKVVAPGLELSAVTPDSTLFNHYTILS*
300	142	390	MQVKFILKYYISFLWKTVTANGETVNMSLLY IFTTMEMRKKSEVGLHLPISILKPFFTIVLDEKI VTGQVWGGELFLLFCKD*
301	97	273	MPELPTWVLALLPHPVVLLIDSGELEAFEQIC RSTLKAVWHSVHGAMSVCFICFTFCH*
302	34	279	MRIVRRMCMWSAGPAPATVCAVMVAAPKSP QSPPRWACVYSLIGCHSSDPFSVYFSGISWRDI SLSLYSMAQESQNQSILK*
303	159	407	MCPLLVYKIILVFAAMFFFSQGSQVEIRSHEG EHCVGTVHLLSHFLYSKNNPVFYKGNTSFIFE TMEEDSLSSLAERSGSCM*
304	347	700	MPQFPVAFGIMFTYFTLAHKVLHSQASACLFS IICFFPTCTLHFSQVGSHAAPWMGHDALCLRV FLYRLPCEKPSPSAHMVTGSVLEGPLCALALS SFPPGATLHLSCLSLKRAVFFY
305	350	568	MSWRTRSMHTHISVSFKGKIRPTSAYLLLFLF FFCYGVSLCCPGWSEVVARSRHLASSASRVH AILLPQNPE*
306	110	250	MLSTLSIGTLSMLIIVVSDSWSYSSNSPAMFGS DAGFIPSNCIFAF*
307	625	906	MDPPCPWLHPAAWPLQTPLALPLLGTGSSPM PIFRWRPPVHLLSMAQGPSFLAGAARGDKAK GAPRRHGANFALTRWAYPIRALNLLGGRQT W*
308	145	306	MILVSLLILIVEPLFASLTPLSLCFECVVFLNVG QHLTDQTFSLNGLLFLSNS*
309	19	222	MSPLLPLSYKLVLCFPTPNGVVTHGEQNASST DIEHGLKTILIKPPARILKRKTEGEESNRLTLPT T*
310	273	533	MGPVSGCWHMSLCLRVYLALDPAHQELMPP GSSLQPITLGIGIEILQPPTLEVGNSEALSVPSR RTPRRTELPWPTVLTGFLINTL*
311	13	171	MLTCVPERLFQCHHLIRMTCLFMILEFRLFKY DSNLCSHVIINHPQVQGRQR*
312	97	249	MWGAPALQCIVFFRWTRSKCLPDTGNVCTKT QRKKAAGRLGVAGGIALGL*
313	126	308	MSLRIRAARNWARDVQKLWTIVVLLVLILIRS AVNLLINSRTEDKSLQLVLYQSVIICFP*
314	34	210	MVWHVRKSSFVWLLQLFSFISCHSVISVSPVH

SEQ ID NO:	Predicted beginning nucleotide location corresponding to first amino acid residue of amino acid sequence	Predicted end nucleotide location correspondi ng to first amino acid residue of amino acid sequence	Amino acid segment containing signal peptide (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
			VPWTQCAVIPPYTSCPKLFAIQGGRF*
315	146	388	MGGFKVNLFVKVKAEGSPLCWLKLACLGAW LLSLLISQKSDEQLCFGLLTWNFSTSDSILVWF VGPRAGTQSKPVVSYKF*
316	110	265	METSSAFTNPLLVCFLALLHSVMNITYTPPKK KNENCSKPLILTSSLGTVQ*
317	457	675	MCAFLLPFKLLFFLEISLAMKSHFPFTLLILSR VLLKKTLYVLKLGWLITIPSNDLTSVFTLMIH RQNQKHF*
318	260	51	MPTLLQVMSWMLSFGTQTLQLESCTCALHIV GAWKVPYPLFSRVLICQVKILSTSISQEKVFRT ESRTE*
319	38	175	MCRALLLLLCSPNSSFQWLPLPVHPHTTIRYR SYNMVPVKLTNVQ*
320	77	244	MLFLHIAECSFLRLKVAFPSSLNFQPLAQFLA HILEVFYKCLWKKGVQVFNFLAN*
321	384	211	MPLDTDAILHRTAEWYVLCLITCIFMYVLYVP YLRSLILLEYLHLLPFEILIQANAG*
322	155	304	MATRSKGAFINCYIILLLTFLMIRTFYNLMEY YCPTLLIRKLMSNTKIL*
323	13	171	MLTCVPERLFQCHHLIRMTCLFMILEFRLFKY DSNLCSHVIINHPQVQGRQR*
324	173	397	MDFLSRLMLLRMCKCVTATYQYIRRSLFLNL VPLLQTLSILTAHSVLLRPALSSLVKMEDSQA LSLSLEPESAF*
325	268	450	MFCLWNQWVVTAQRLLVSWLSHAQRQPCPL SLFCGRRNPLAWTIFGWKHQPLTSDCHFQM*
326	63	239	MTTSSLVLPPLFVLKCQRFYPPLYLHPYSICQ HVSILVKIVWTWGSEVPTLGTIEIGT*
327	141	359	MLAWRLLCPWGPGLPTTTARSGERTERRERV RTASPRKILFKTQPPRGSSTDRCPWGRQCLHG TGTCHMPNR*
328	217	357	MTSRPHFFRYLCSLPPLLFPLLXQSQLLPGSPL PIALQSRVGSLLA*
329	99	386	MLPSFLPQSLGNLIHTLGFLLIIHKYMSAFKNR TDEFMNMGMQPYIKSPYRLSMSQISLKFDLS QTDLILPHKFYSPSSFPTVMLFYSFGRLSHKP
330	99	248	MQAWRSFVMGVEVLMYIVAVRCRAVFATSL WQPWCYTRAGGQFNVSQAR*
331	11	244	MVVLYIVRAYNHYILCCLSSSLYLVLILLVTV YLMLTTSSYNDVSLVIWIASSFASSKFFRKGL REYSYFMNFLARS*
332	71	349	MLEWPLLGQILPMIIPLPPLPALVVWPIGLTHC

SEQ ID NO:	Predicted beginning nucleotide location corresponding to first amino acid residue of amino acid sequence	Predicted end nucleotide location correspondi ng to first amino acid residue of amino acid sequence	Amino acid segment containing signal peptide (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
			PWPSPFVPASLDGFYNSRSLDGKSPPLRPEKW SPWSWFSRLPGHGLPKEGGQRKREVWA*
333	163	360	MMLWQVYPGPSAAVLCLFLHPPWSRSTAVE REKRQKDGRGQRMLLPQPQCLMSSCCLVDV QSLTG*
334	61	243	MFFCLTRQSLLCTLLLMLKRCIFFSYCVICRA KSFELFTSEITFPDKRAKQCLFKLFSGT*
335	229	387	MGNTDILLLLSLFCFSYELVAGKTKAQFGVPF AEFSVFLILENVTACRFLYI*
336	146	352	MSILVVSAFLANLWLLMTISTSQMLNMTKIT YLVLFLHLSALRIGSTPHSFLLKSYHLGTHFSL FHMNS
337	308	72	MFVFVSIHTELVPILRPLCLLYCCPDCSVPRPL YSLKYLLLINDFPELQIHMSSFSQSLHYIILSYF FHSICHILLVL*
338	362	204	MMRVIILIWFRISKGTFQHSTTKCDVCFRVFL LSNCSFLSLNYKLTSDFIIY*
339	88	276	MQMVVPRLLSVPQLLNTAPLFLPWEKTVKTQ YSGIIFKFKSRIETAEKSIGDTKERIQPSQI*
340	103	282	MLLPVFLLYLSQDLADSRAPAHCSVNTDLHL KWGSLCVLSHFQVDLPVNPICEHICRCP*
341	53	190	MNLEHVIVSLLTFYRVLLYKEIIGLHHCFQHF HVNAFLLSPLPPS*
342	145	2	MSALWLVYVSCGCSCALCVLVHAQTCTSFSI LVHAQAHVLSCDHIHTS
343	184	336	MLIQFTFFFPISQRFWFCLLFFFPQTFKCMKFY SLIEVREGVCIHQRSQFV
344	295	17	MIELAWKFIMHINALLSFGTTSLKIKFAHVFSF LQLCIMNKITLISLNTPPFQEEFTELGSFVNHC ELINKNLSHNLPFFPWLIILLKVIRY*
345	81	290	MPNRIELSALFRVPFHVDWLLFAFVFFPLCYV LGFCLPSTVKTHSLIVGAVFQSEIVTNVCQIYK CKTT*
346	867	1091	MRAHPHTPRSAAHIILSTWLALGPYLYRTRNC GPLTEEPSLAPSTQGARRTNTTPELAARPVAP CLHISYTFTRG
347	101	316	MPLLYIICLRQLVLFHSKCHSQHSCRAGGIQY SMHVSLFLSSPINYDNGFLVSPTFPLHVKLSFL KYSFKCI
348	823	1002	MGPKRGLFFFIFFLDTEPSVLGGGGGGQYGLT RTHLWRQGASYLTLLRNGTQSGPLAHL*

Table 7

SEQ ID NO:	ACCESSION NO.	SMITH-WATERMAN SCORE	% IDENTITY	DESCRIPTION
349	M84913	59	36	Antirrhinum majus DEL
350	U11271	110	71	Homo sapiens thromboxane A2 receptor
351	AE003824	696	54	Drosophila melanogaster CG13189 gene product
352	X85236	81	29	Trimorphomyces papilionaceus cob
353	AF068065	166	29	Cryptosporidium parvum GP900; mucin-like glycoprotein

Table 8

SEQ ID NO:	ACCESSION NO:	DESCRIPTION	P-VALUE	RAW SCORE	RESIDUE POSITION
352	BL01253	Type I fibronectin domain proteins	6.894e-07	14.35	55-94
353	PF00624	Flocculin repeat proteins	8.893e-09	13.62	53-108

Table 9

SEQ ID NO:	SIGNAL PEPTIDE POSITION IN AMINO ACID SEQUENCE	maxS (MAXIMUM SCORE)	meanS (MEAN SCORE)
352	1-18	0.971	0.925

5

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-87, 175-261 or 349-353, a mature protein coding portion of SEQ ID NO: 1-87, 175-261 or 349-353, an active domain coding portion of SEQ ID NO: 1-87, 175-261 or 349-353, and complementary sequences thereof.

2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.

3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.

4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.

5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.

6. A vector comprising the polynucleotide of claim 1.

7. An expression vector comprising the polynucleotide of claim 1.

8. A host cell genetically engineered to comprise the polynucleotide of claim 1.

9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.

10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:

(a) a polypeptide encoded by any one of the polynucleotides of claim 1;

(b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO: 1-87, 175-261 and 349-353; and

(c) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 88-174, 262-348 and 354-358; the mature protein portion thereof, or the active domain thereof.

11. A composition comprising the polypeptide of claim 10 and a carrier.

12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and

b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.

14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;

b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and

c) detecting said product and thereby the polynucleotide of claim 1 in the sample.

15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.

16. A method for detecting the polypeptide of claim 10 in a sample, comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and

b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.

17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and

b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

19. A method of producing the polypeptide of claim 10, comprising:

a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-87, 175-261 and 349-353, a mature protein coding portion of SEQ ID NO: 1-87, 175-261 or 349-353,

an active domain of SEQ ID NO: 1-87, 175-261 or 349-353, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-87, 175-261 or 349-353, under conditions sufficient to express the polypeptide in said cell; and

b) isolating the polypeptide from the cell culture or cells of step (a).

20. The isolated polypeptide of claim 10 comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 88-174, 262-348 or 354-358, the mature protein portion thereof, or the active domain thereof.

21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.

22. A collection of polynucleotides, wherein the collection comprising the sequence information of at least one of SEQ ID NO: 1-87, 175-261 and 349-353.

23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.

24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.

25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.

26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

29. A method of detecting bone marrow cells or tissues in a sample comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form a complex; and

b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected

wherein the presence of the polynucleotide of claim 1 indicates the presence of bone marrow cells or tissues.

30. A method for detecting bone marrow cells or tissue in a sample comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form a complex; and

b) detecting formation of the complex so that if a complex is detected, the polypeptide of claim 10 is detected,

wherein the presence of the polypeptide of claim 10 indicates the presence of bone marrow cells or tissues in a sample.

SEQUENCE LISTING

<110> Hyseq, Inc.
John E. Ford
Bryan J. Boyle
Y. Tom Tang
Chenghua Liu
Vinod Asundi
Ping Zhou
Jie Zhang
Radoje T. Drmanac

<120> NOVEL BONE MARROW NUCLEIC ACIDS AND POLYPEPTIDES

<130> 21272-019

<140> Not Yet Assigned
<141> 2001-01-25

<150> 60/250,583
<151> 2000-11-30

<150> 09/663,870
<151> 2000-09-15

<150> 09/631,451
<151> 2000-08-03

<150> 09/617,746
<151> 2000-07-17

<150> 09/491,404
<151> 2000-01-25

<210> 1
<211> 1106
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(1106)
<223> n = a,t,c or g

<400> 1
gattggtcca tttcatcaaa gttgctgaat ttatgtgcat agagttatgc ataatattcc 60
ccgtagtgat gtcttctcct tcattcttca tattagtgtc tgaaaacaat aacaatgatt 120
tatttctcat tctgatggtt cttctcctgg tctctcttgg gcttactctg gctgtattgg 180
tagattggct acagctggag gatctaaaaa tggcctcacc catatgtttg gggccttgat 240
ggggatgact ggagtagctg agacttgtct ctccctggat tcattgggaa ccaaaatgat 300
ctgccatgtg gaagtggaag cataaggaga ggacttgatc atagtaaaat cctcagtaaa 360
tgctcatgga tttatccaat gacgtggaga ccaaagaaag agtttcaaag aggaaggggt 420
atcagtatgt cacatgctgc aaaatagtta agcagaatat ggcctgagaa ggctattgtt 480
ttgttttggg ttagtaactt ttgagagagc atttttagta gagtttagaa aagattgcag 540
agcattaaga agaaaacgtg gtgaagaaaa ggaggcagtc tgtatttttg ttctgagaaa 600
ttttatttag aaaataaatg ctagttaaat ggaacaggag actaaacaga gaattgtgag 660
agacccaatg catatttgaa ggtaaattct gtaaagaagg gaaagcttga aaggggtcaa 720
gtaggggggaa ccatgggagg gggccagaag taaaatccct ggaaagagct ctgggagagg 780
atgtgataaa gagcgcagtt ggagggattg tttggtcttg gaaggagctg gctagcctga 840
aaactggaag acagaagtac gggaaggcga agaaaagaat agagaagata gggaaattag 900
aagatannnn nnnnnnnnnn nnnnnnaaaa gataaattta aacctgaaaa gtaggaagca 960

gaagaaaaaa gacaagctag gaaacaaaaa gctaagggca aaatgtacaa acttagaaga 1020
aaattggaag atagaaacnn nnnnnnnnnn nnnnnntaag gcggcggggg ccggaatttc 1080
caaatttata attggaaaat ccggag 1106

<210> 2
<211> 1586
<212> DNA
<213> Homo sapiens

<400> 2
tttcgtggat ctgataaatg cctgtagtca ttatggctta atttatccat gggttcacgt 60
cgtaatatca tctgattctt tagcctgata atgaattata cagaagatct ttcaaaatta 120
cagagctctt atatggagga ccttcatttg acatagcttc cattattccg atcttggagc 180
cactttcaga agacactatt gccggcctca gtgtgcatgt tctgtgtcgt acacgcttga 240
aagagtatga acagtgcata gacatactgt tagagagatg cccggaggca gtcattccat 300
atgctaatca tgaactgaaa gaagagaacc ggatagactc tgtggtggaa aaaactgttg 360
cctgaacttt gtcagagaat aaaatgtggt ggagagaagt atcaactcta cctgtcatca 420
ttagaagaaa cattgtcaat tggtgctgtg gaactagaac tgaaggaatt cttgaatgtt 480
ctcccaaaag atggtcctgc cacatttttc ttgccatatc ttctctattg cagtcgaaag 540
aaaccattga cttaaaggta tcatttgaaa aataccataa tggcatttga gactgaattt 600
ctaaaaattg aatgccaaag tacaagtaga ggagtttttt attttatata tcacacacac 660
acacacacac acacacacac acacacacat atatgataca aatgctttca ggctgcttac 720
cttaccgtgt agtggtaact attcacttct taatttatga cctcaatcaa tttaattgtc 780
tagaatgtaa aaagtcttta agacataaga attcctcaaa gaagccatac atttttaag 840
ggtggggatt gacttttatt ccaaggaaca acatcagttc actgttgttg gagacatgac 900
aatcattttc atcccaagaa cactttaagg aaacattttt acaagtatgc ttgaaagaat 960
gtcaactaac tggtccagaa ttttatcttc ttgatttttc cagatttctc tatgtttttt 1020
gaggaagaat gttaatgttt tgccatggta aaagatttca aaacctcatt ttttttgttc 1080
ctttttcttg ttacttttta ggaaaaactc atgctctgtt tcttctgaat caaatgaagt 1140
agaagtttac aaagctaact ttcttcttgt ctagctatta acatgatttg tcaacatgca 1200
tgttttcttc agccaaagcc ttggtttcca ttttttgttg atgtgtacat cttgctcttt 1260
tagctagagt gtatgtgaaa ataaagaaat acatcattgg taatcacaac catgtggctt 1320
caattataac tttttgttta aaaaattttt agttcaagtt tagttcattg atattatcct 1380
ctgaatgcag ttaaggctgg gcagaaattc tactcatgtg acatctgcca caggtctatt 1440
ttgaagcttt tcttctaatg gcaatgtttg tccttaccag gatttaatct atagaattgt 1500
ctctcaactc tgcttttctc cagttccaga taacgtcctt aagaccatct gttcagggga 1560
tcacaaaact caaacccacg cgtccg 1586

<210> 3
<211> 4404
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(4404)
<223> n = a,t,c or g

<400> 3
atggcgtcat ctaatgtaaa tgatataatg ccagtcccag aagccttggt gtatgtgggt 60
ccccgaccta ctgaaaacag atgtaatgtc tggagcatta gcagccacct tggatcagta 120
gatgaccttg agaatgaact cacagtacag ctgagaaatt acctggaaaa ggctggaaat 180
tgtataagta cctggaaaag actaaggtgt gctggcccca tgaccgtcaa ggcggagaaa 240
ggtgcggtcc aacccgcgga ggctcgaccc ggaagccgcc gtggccgccg cacaactgcg 300
ctcgtctctc ggctggggag ctccgtgtcg caccgggttc ttgggctggc cgtgtccttc 360
tccctcggtc gccactggcc catgagtccc catggcgacg gcaggggcca ggcccagggg 420
cgggcagtcc gagtggggct gcggcgctcc gggacaatcc gcggcggcgt ggcggtgttc 480
gccgccgtgg ccgcagtgtt caccttcacc ctgccccctt cggtaccggg gggagactcc 540

ggggaactga tcacagccgc acatgagctt ggagttgccc atcctcctgg ctatcctttg 600
ttcacgctgg tggctaaact ggcaattaca ctgtttcctt ttggttcaat tgcctaccgc 660
gtcaatcttc tctgtggctt atttggagca gtagctgcnn cncgactttt tttcaccgtt 720
ttcaggtagc taaaattggt gctttctgct gtggccttag tttatgtaac cagcacacaa 780
taatactcta tgttttgtgc ataatacctt ggattctctt tcaactttta aaaaagaagg 840
aactctccct gggctctttg ttgaagttga gcctgtactt ctctgctggt ttgctgccct 900
atgtccacct tcccatctca tcttacctta atcacgcccg gtggacctgg ggagaccaga 960
caacactgca aggatttttg acacattttc tcagggaaga atatggaacc ttcagcctgg 1020
ccaaatctga aataggatcc agtatgtctg aaatactgct ttctcaagta acaaatatga 1080
ggaccgaact ctcattcaac atccaagccc ttgcagtttg tgcaaatata tgtttagcaa 1140
caaaggacag acagaatcca tcattagtat ggctttttac tggaatgttt tgcatttatt 1200
cattgttctt tgcttggaga gcaaatttag atatttcaaa accacttttc atgggtgtgg 1260
tggaacgatt ctggatgcag agcaatgcag tagtggccgt ccccgctggc attggtttgg 1320
ctgcagttgt gtctgagact aaccgagtgc tgaatagcaa tgggcttcag tgtctggaat 1380
ggctttctgc aactcttttt gtagtttacc aaatatattc taattacacg cgcttgtgac 1440
caaaggacca actatgtgat cgataagtac gcaaagaacc ttctcacctc tatgcctcat 1500
gatgcaatta tcttactcag aggagatatg ccaggaaatt ctctccgtta catgcattac 1560
tgagaggggt tgaggcctga catttcatta gtggatcacg aaatgatgac ttacgagtgg 1620
tatttaccca acatggctaa gcacttgcca ggtggcaact ttcctgggaa ccggaggaat 1680
cctgtggaag gaatattacc tagtggaatg gtcacattta atctttatca ttttcttgaa 1740
gtaaataaac aaaaagaaac atttgtttgc ataggaattc atgaaggcga cccaacctgg 1800
aaaaagaact attcactttg gccatggggg tcttgtgaca aattagttcc tttggagatt 1860
gtattcaacc ctgaggaatg gattaaactt acaaaaagta tctataactg gaccgaagaa 1920
tatggaaggt ttgatccatc ttcttgggaa tctgtggcca atgaagaaat gtggcaagcg 1980
aggatgaaaa caccgttctt catctttaac ctggcagaaa ctgctcacat gccttcaaaa 2040
gtgaaagctc aactctacgc tcaagcatat gacctttata aggagattgt ctatttacaa 2100
aaggagcacc cagtgaattg gcacaagaac tatgccatcg cctgtgagcg gatgctgcgt 2160
cttcaggcaa gagatgcaga tcctgaagtg ctgttatcgg aaaccatcag acatttccgt 2220
ctgtactctc agaaagcacc gaatgaccca cagcaagctg atattttagg tgctctaaag 2280
cacctaagaa aagaactgca aagtctgaga aataggaaaa atgtctgaga cagcaaaata 2340
tgaaaaacct gctcatcgtt cagcttccaa aattctgaag tctggaagtt tttccttcaa 2400
agaaaagaaa ctgcataaaa aatttaaaac taagtcatct cccagatata agtatcatgg 2460
tccagcagta ctgtttaatg gggtattcag tgactaaagg tctgctattt atgcaaaatt 2520
ctgtttatcc cgtgttacca aaattaccat ttcagtgaga agcttttgaa aagtcttctg 2580
acttccagtc tttcaccaga tgactgcact ggattagatt ctagaagaga atgaaccatt 2640
ttcatataac taaatattgg tcatgaactg tgtaaggggc catgcttatt ggggtccagt 2700
tttaaagtta aattcttttg atattaatac cagaccaaag acattttctg tttcctggaa 2760
aaaaaaaaaa atgaatcatg ttaggcttta ggtgagagta catttttttac aaagtagcta 2820
tagttgttac atagtcttac acttcaagct aaacaccaaa tgggtgatat tttgaaaaaa 2880
gtttgtgttt tactgtctta gatcgttctt ggaaatcact aaaaaaaaaa aaaagttaat 2940
ttgatgtttg cttatttcag ttgcagaaac tggcgagtaa aaaagatttt gcatttactt 3000
aattaatttt atatttatgt tttatttcta tttggactca gagatctaga cccaataatt 3060
agtaggctcc tgctgccaaa atgcaaaggg gaaaaaaagg cagggattgg gggggatggg 3120
gaagtggccc aataaattca tttctgttac aagcatcttt tatacatatt actaaagaga 3180
acatagtaag aaatgcaaat aatagttctt tattttatta tgacagttaa ttaatagcac 3240
cctgttttaa tgaataaagt cactcagagt ccttcagcac ttcctaagta gaggccagaa 3300
tctgtagggc ttcttttccc cccaattgta tagctcctag actccaagca ctatataggc 3360
ccctgtatag aaatgctcac taatgaagag ggagggctag aagcttgtct gcattcaaag 3420
atcactggtg agtcattcag caagaaaagg ccccttacca gggaatagtc acagttccgt 3480
gggcattgta ctaggcaaaa gggtctgatc aaaggtctcc tgtgggagct tgcatgggtt 3540
ccctttttcat actacggacc ataattaaaa ccactaattc tctttttaaa atgctggcgg 3600
ggtggccatg tagggcatct tgtctgggag tgtcccttgt ggatgtcata aggctgttaa 3660
gggaccagtg ccgagggctt ttttgagtga catagctcag tctgaagggg tcttaaagac 3720
cacgggtgct ctctaaaaga cttgacaggg cgcttgactg cacaattcga gctgaatttg 3780
ccccttttgtc agctgccagt aaataaatct caaaggggga aaagctgaag tttcattacc 3840
tgatccatgg ggctttgttg gttttggcat cacacagggg gaagctcttg cccctccatt 3900
ctctggattt gaagatgtcc attggagcct tgcagtgcct ggaacaggtt tcagagcgga 3960
acctttttgaa gagtgtcaat agttgtaaca gttccagctg ttaggaagac caaataaatg 4020
gaggagctca ttaatccgct tttggctcct cagtgccctt ttgccccttt tatcacaggc 4080
ctttattagg ctccgtactc atctttgaac caggaaaaaa ttgaattgaa gttgttggag 4140
aactaattgg caaggatttt taatcatggg cccaagaact ttcactgact tgaaagtaac 4200
ttctccacag ggaaggacca aaaacctggt ttaccttaaa acgaaaaacc tgttggagtt 4260
caagcgtggt gtaaaaatgt aaggaagcat tgataaattg tctaagttta tccatttgaa 4320
agaaattgtg taagattatg atattctctt ttctttaaaa aaaaaagtac aataaaattc 4380

aaacattcct taggaaaaaa aaaa 4404

<210> 4
<211> 431
<212> DNA
<213> Homo sapiens

<400> 4
tttcgtggaa cctgtgatcg agttacataa agacacacgt aatcagaact gagcaaagct 60
cttctctaat atgtaggtga aaaaattaac tgactctgtt aatctggaaa atgattgaat 120
gtgtatttgt taaattttaa tggaaatatc caaaaaacca actcagacaa gtggtgaaca 180
gagtagccgt ggagtattta tttcttctgt atatgtatat tgcaattgag tgcagaattt 240
gggatacatg gattcctttc cttcttcctt tgtgttctct ccccagtatg ggaagactct 300
ctatgataac taccagcggg ctgtggcaaa ggctggactt gcctctcggt ggacaaactt 360
ggggacagtc aatgcagctg caccagctca gccatctaca ggacctaaaa caggcaccac 420
ccagcccaat g 431

<210> 5
<211> 1463
<212> DNA
<213> Homo sapiens

<400> 5
atggcatcga tatcagagcc tgttacattc agagagttct gcccgttgta ctatctcctc 60
aatgccattc cgacaaagat ccagaagggt ttccgctcta tcgtggtcta tctcacggcc 120
ctcgacacca acggggacta catcgcggtg ggcagcagca tcggcatgct ctatctgtac 180
tgccggcacc tcaaccagat gaggaagtac aactttgagg agtttaatag tctgaaatct 240
ctcaaatgct tcagacactt tataaggaag actcgtgaat ccaacgatga accccaacaa 300
gatctatcaa agaaatccat gcacagaggc accctagtca cactgctaaa gatcacaatc 360
ttaaaatccg tgcacagagg cgtcctagtc acactgctaa acatcacaat cttaaaatcc 420
atgcacagag gcgtcctagt cacactgctg aacatcacaa tcttaaaatc catgcacaga 480
ggcgtcctag tcacactgct aaagatcgct gtcttaaaat ccatgcacag aggcgtccta 540
gtcacactgc taaagatcac agtcttaaaa tccatgcaca gaggcatcct agtcacactg 600
ctgaagatca ccatcttaaa atccatgcac agaggcgtcc tggtcacact gctaaagatc 660
accatcttaa aatccatgca cagaggcgtc ctggtcacac tgctgaagat cacaatctta 720
aaatccatgc acagaggcac cctagtcaca ctgctgaaga tcacagtctt aaaatccatg 780
cacagaggca ccctagtcac actgctgaag atcaccatct taaaatccat gcacagaggc 840
atcttagtca cactgctgaa gatcaccatc ttaaaatcca tgcacagagg catcttagtc 900
acactgctga agatcaccat cttaaaatcc atgcacagag gcgtcctagt cacactgcta 960
aagatcacca tcttaaaatc catgcacaga ggcgtcctag tcacactgct gaagatcaca 1020
gtcttaaaat ccatgcacag aggcgtccta gtcacactgc tgaagatcac catcttaaaa 1080
tccatgcaca gaggcaccct agtcacactg ctgaagatca ccatcttaaa atccatgcac 1140
agaggcgtcc tagtcacact gctgaagatc accatcttaa aatccgtgca cagaggcgtc 1200
ctagtcacac tgctgaagat cacagtctta aaatccatgc acagaggcac cctagtcaca 1260
ctgctgaaga tcaccatctt aaaatccatg cacagaggcg tcctggtcac actgctgaag 1320
atcacagtct taaaatccat gcacagaggc accctagtca cactgctgaa gatcaccatc 1380
ttaaaatcca tgcacagagg cgtcctagtc acactgctaa agatcacaat cttaaaaata 1440
gccagagaac aatgacgtat tac 1463

<210> 6
<211> 3796
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<222> (1)...(3796)
<223> n = a,t,c or g

<400> 6

ctcctgcata gagggtacca ttctgcgctg ctgcaagtta cggaatgaaa aattagaaca 60
acagaaacgt ggttctcttg acacttcagt gttagggaac atcagcaaga cccatcccag 120
gagaccttga aggaagcctt tgaaagggag aatgaaggag tcatctttgc aaaatagctc 180
ctgcagcctg ggaaaggaga ctaaaaagga aaacatgttc cttcagtcgt caatgctgac 240
ctgcattttc ctgctaatat ctggttcctg tgagttatgc gccgaagaaa atttttctag 300
aagctatcct tgtgatgaga aaaagcaaaa tgactcagtt attgcagagt gcagcaatcg 360
tcgactacag gaaattcccc caaacggggg gcaaatattg gacagatcta gaacttgctg 420
ataatttcat tacacacata accaaagaga tcatttcagg gcctggcaaa atctcactaa 480
ataaaatcta aacccaaacc ccaatgtcca gcaccatgaa cggatatccg ggtatacaat 540
caaatggctt gaatatcaca tacggtgcat tcctcaacct aaaaaaccta agggagttac 600
tgcttgaata caaccagtta ccccaaatac cctctggttt gccagagtct ttgacagaac 660
ttatgctaat tcaaaacaat atatacaaca taactaaaga gggcatttca agacttataa 720
acttgaaaaa tctctatttg gcctggaact gctattttaa caaagtttgc gagaaaacta 780
acatagaaga tggagtattt gaaacgctga caaatttgga gttgctatca ctatctttca 840
attctctttc acacgtgcca cccaaacttg ccaaagctcc gtgctgcaaa ctatttctga 900
gcaacactca gatcaaatac attagtgaag aagatttcaa gggattgata aatttaacat 960
tactagattt aagcgggaac tgtccgaggt gcttcaatgc cccatttcca tgcgtgcctt 1020
gtgatggtgg tgcttcaatt aatatagatc gttttgcttt tcaaaacttg acccaacttc 1080
gatacctaaa cctctctagc acttccctca ggaagattaa tgctgcctgg tttaaaaata 1140
tgcctcatct gaaggtgctg gatcttgaat tcaactattt agtgggagaa atagcctctg 1200
gggcattttt aacgatgctg ccccgcttag aaatacttga cttgtctttt aactatataa 1260
aggggagtta tccacagcat attaatattt ccagaaactt ctctaaactt ttgtctctac 1320
gggcattgca tttaagaggt tatgtgttcc aggaactcag agaagatgat ttccagcccc 1380
tgatgcagct tccaaactta tcgactatca acttgggtat taattttatt aagcaaatcg 1440
atttcaaact tttccaaaat ttctccaatc tggaaattat ttacttgtca gaaaacagaa 1500
tatcaccgtt ggtaaaagat acccggcaga gttatgcaaa tagttcctct tttcaacgtc 1560
atatccggaa acgacgctca acagattttg agtttgaccc acattcgaac ttttatcatt 1620
tcacccgtcc tttaataaag ccacaatgtg ctgcttatgg aaaagcctta gatttaagcc 1680
tcaacagtat tttcttcatt gggccaaacc aatttgaaaa tcttcctgac attgcctgtt 1740
taaatctgtc tgcaaatagc aatgctcaag tgttaagtgg aactgaattt tcagccattc 1800
ctcatgtcaa atatttggat ttgacaaaca atagactaga ctttgataat gctagtgctc 1860
ttactgaatt gtccgacttg gaagttctag atctcagcta taattcacac tatttcagaa 1920
tagcagggcg taacacatca tctagaattt attcaaaatt tcacaaatct aaaagtttta 1980
aacttgagcc acaacaacat ttatacttta acagataagt ataacctgga aagcaagtcc 2040
ctggtagaat tagttttcag tggcaatcgc cttgacattt tgtggaatga tgatgacaac 2100
aggtatatct ccattttcaa aggtctcaag aatctgacac gtctggattt atcccttaat 2160
aggctgaagc acatcccaaa tgaagcattc cttaatttgc cagcgagtct cactgaacta 2220
catataaatg ataatatgtt aaagtttttt aactggacat tactccagca gttccctcgt 2280
ctcgagttgc ttgacttacg tggaaacaaa ctactctttt taactgatag cctatctgac 2340
tttacatctt cccttcggac actgctgctg agtcataaca ggatttccca cctaccctct 2400
ggctttcttt ctgaagtgca gtagtctgaa gcaccgcgat gtaagttgca atcggctcca 2460
aacaatcaac aaatccgtgc ttgaaactaa gaccaccacc aaattatcta tgttggaact 2520
acacggaaac ccctttgaat gcacctgtga cattggagat ttccgaagat ggatggatga 2580
acatctgaat gtcaaaattc ccagactggt agatgtcatt tgtgccagtc ctggggatca 2640
aagagggaag agtattgtga gtctggagct aacaacttgt gtttcagatg tcactgcagt 2700
gatattattt ttcttcacgt tctttatcac caccatggtt atgtgtggtg gtgtgtgtgt 2760
caccatttgt tttacttgga tgtttggttt atatataatg tgtgtttagc taaggtaaaa 2820
ggctacaggt tttttttttcc aattccaaac tttttatgat gcttactttt cttatgacac 2880
caaagatgcc tctgttactg actgggtgat aaatgagctg cgctaccacc ttgaagagag 2940
ccgagacaaa aacgttctcc tttgtctaga ggagagggat tgggacccgg gattggccat 3000
catcgacaac ctcatgcaga gcatcaacca aagcaagaaa acagtatttg ttttaaccaa 3060
aaaattatgc aaaaagctgg aactttaaaa cagcttttta cttggctttg cagaggctaa 3120
tggatgagaa catggatgtg aatatattta tcctgctgga gccagagtta cagcattctc 3180
cagtatttga ggctacgggc ccgcggatcc tgtaaggagc tccattcctc cagtggccct 3240
tgacaacccg aaggcagaag gcttgttttg gcaaactctg agaaatgtgg tcttgactga 3300
aaatgattca cggtataaca atatgtatgt cgattccatt aagcaatact aactgacgtt 3360
aagtcatgat ttcgcgccag taataaagat gcaaaggaat gacatttcnn nnnnnnnnnn 3420
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3480
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnng 3540

tctcaaaagc tggcatggta ggtgttcacc acagacatag gcatcactgg ggtcacactc 3600
atggtggtgt tttctggatt caatccctcc tgggctattg gcccaagggc tatactcatg 3660
taagccatgc gagcctctcc cacaaggcag cttgcttcat cagagctagc agaaaagaga 3720
ggttgctagc aagatgaagt cacaatcttt tgtaatcgaa tcaaacaagt gatatctcat 3780
cactttggcc atattc 3796

<210> 7
<211> 3645
<212> DNA
<213> Homo sapiens

<400> 7
ctcaggctca taccacacgg ggggtgctag tgggctgctg tgtgtttaca tgagtaatgg 60
agctgccggg ggaggggagt tgtaagcaga gcgctgagcc tcgcagctcg cattcggagg 120
gaagctgaca tccacaccat gtcgagactt ccagggatgt ggccggggag cagtcacatg 180
ctgtagcttt catgagcaca ggcatcagtc aggcagatgt ttgtcgactg gaatggtttt 240
cagcgccaaa tcttaaaggc agaccacgca aaaagaaacc atgcccacaa agaagagatt 300
cattcagtgg tgttaaggat tcctacaaca catttcgatg gcaaagccag ttgccaaggt 360
gaaatgtgaa gccaggtcag ccttgaccaa gccgaacaat aaccataact gtaaaaaaga 420
ctcagatgaa gaaaaaccaa aggttgccat tggtgaagag cgcacggcag atgaacaagc 480
cttcttggtg gcactttata aatactatga aagaaaggaa aacgcggata gaacgaatac 540
cctatttagg ttttaaccag attaaccttt ggactatgtt tcaagctgct caaaaactgg 600
gaggatatga aacaataaca gcccgccgtc agtggaaaca tatttatgat gaattaggcg 660
gtaatcctgg gagcaccagc gctgccactt gtacccgcag acattatgaa agattaatcc 720
taccatatga aagatttatt aaaggagaag aagataagtc ccctgcctcc aatcaaacct 780
cggaaacagg agaacagttc acaggaaaat gagaacaaaa caaaagtatt tggaaccaaa 840
cgcatcaaac atgaaatacc taaaagcaag aaagaaaaag aaaatgcccc aaagccccag 900
gatgcagcag aggtttcatc agagcaagaa aaagaacaag agactttaat aagccagaaa 960
agcatccctg agcctctccc agcagcagac atgaagaaaa aaatagaagg gtatcaggaa 1020
ttttcagcga agcccctggc atccagagta gacccagaga aggacaacga aacagaccaa 1080
ggttccaaca ctgcagaagc tggcagagga ggcgggagag aaggggccca cacctccact 1140
cccaagtgct cctctggccc cagaaaaaga ttcagccttg gtccctgggg ccagcaaaca 1200
gccactcacc tctcctagtg ccctggtgga ctcaaaacaa gaatccaaac tgtgctgttt 1260
tacagagagc cctgaaagtg aaccccaaga agcatccttc cccaggcttc cccaccacac 1320
aggccaccgc tggcaaacca gaatgagacg gaggatgaca aactgcccgc catggcagat 1380
tacattgcca actgcaccgt gaaggtggac cagctgggca gtgacgacat ccacaatgcg 1440
ctcaagcaga ccccaaaggt ccttgtggtc cagtcgtttg acatgttcaa agacaaagac 1500
ctgactgggc ccatgaacga gaaccatgga cttaattaca cgcccctgct ctactctagg 1560
ggcaacccag gcatcatgtc cccactggcc aagaaaaagc ttttgtccca agtgagtggg 1620
gccagcctct ccagcagcta cccttatggc tccccacccc ctttgatcag caaaaagaaa 1680
ctgattgcta gggatgactt gtgttccagt ttgtcccaga cccaccatgg ccaaagcact 1740
gaccatatgg cggtcagccg gccatcagtg attcagcacg tccagagttt cagaagcaag 1800
ccctcggaag agagaaagac catcaatgac atctttaagc atgagaaact gagtcgatca 1860
gatccccacc gctgcagctt ctccaagcat caccttaacc cccttgctga ctcctacgtc 1920
ctgaagcaag aaattcagga gggcaaggat aaactcttag agaaaagggc cctcccccat 1980
tcccacatgc ctagcttcct ggctgacttc tactcgtccc ctcatctcca tagcctctac 2040
agacacaccg agcaccatct tcataatgaa cagacatcca aatacccttc cagggacatg 2100
tacagggaat cggaaaacag ttcttttcct tcccacagac accaagaaaa gctccatgta 2160
aattatctca cgtccctgca cctgcaagac aaaaagtcgg cggcagcaga agcccctacg 2220
gatgatcagc ctacagatct gagccttccc aagaacccgc acaaacctac cggcaaggtc 2280
ctgggcctgg ctcattccac cacagggccc caggagagca aaggcatctc ccagttccag 2340
gtcttaggca gccagagtcg agactgtcac cccaaagcct gtcgggtatc acccatgacc 2400
atgtcaggcc ctaaaaaata ccctgaatcg ctttcaagat caggaaaacc tcaccatgtg 2460
agactggaga atttcaggaa gatggaaggc atggtccacc caatcctgca ccggaaaatg 2520
agcccgcaga acattggggc ggcgcggccg atcaagcgca gcctggagga tttggacctt 2580
gtgattgcag ggaaaaaggc ccgggcagtg tctcccttag acccatccaa ggaggtctct 2640
gggaaggaga aggcctctga gcaggagagt gaaggcagca aagcagcgca cggtctgtat 2700
tccgggagcc tgtgtaactc gggcctcaac tccaggctcc cggctgggta ttctcattct 2760
ctgcagtact tgaaaaacca gactgtgctt tctccactca tgcagccccc ggctttccac 2820
tcgcttgtga tgcaaagagg aatttttaca tcaccgacaa attctcagca gctgtacaga 2880
cacttggctg cggctacacc tgtaggaagt tcatatgggg acctttttgca taacagcatt 2940

```
tacccttag ctgctataaa tcctcaagct gcctttccat cttcccagcg tgtcatccgt   3000
gcaccccagt acaaaactgt aggctcagct ctgcccagca gtctaaagcg gcatggccaa   3060
cagagcttca ctccttaccc aggagtgctg gcttatagag ttagaagtca gtatttcttc   3120
taatctgagg ctatgatcag tccccagctg taggggccca gaggggaggt gaacatgcct   3180
gatttttgtg ggacaactgt agcccacaaa ctgactggct ggtgagtctt gactcccttc   3240
caacacagat gcccaggcac ctccagatca ttcacttcgc acgtgggcct tgtgaaggga   3300
tttgtgaata tccaggaaga acttagagga ccccatctga gttcggatgg tcaggaaaca   3360
atctgggcaa aaaagaggca ggcatttcaa aggaaggggc aaggaagact ggcaaacaga   3420
tggcaaggga tgcccctctt tttcataaaa ctctccaagg ttcaatcaat gcaatgtata   3480
gtgaaacttc aatagatctt tcattttgac actattaaac aatccagaga agtaaacact   3540
gttaaattga ctgtatatat tcgcttctta caactacctg tatcactgct tgctcaccta   3600
atttatatac aggtagttcc attttctccc agttccttct cgtct                   3645
```

<210> 8
<211> 870
<212> DNA
<213> Homo sapiens

<400> 8

```
tttcgtggaa gcaggaagtg actgcgggag tggagccggc gagagagtgg cagcgggggc     60
tgatggaagt gcagtggggg ctggagaggg caccctagtg agttggcttt tctctccttc    120
ggtgcctttg gttgggctcg aggcccgggg tcggaggtta tcagttactt gagtgcggga    180
cttggaaggg caggcacctg gagcttggaa ggggctgtat ccagcatgct ccaaggccac    240
agctctgtgt tccatggcct tgctgtggac cttcttcacc ctgggggatg acagcagctg    300
gggcagctct cgtgttcgta ttctctagtg gacagagagc ggatcttaag atggaagtct    360
tgtctttgct gcaggggtca tgttggcagc ttcctattgg tctcttctgg ccccagcagt    420
tgagatggcc acgtcctctg ggggcttcgg tgcctttgcc ttcttccctg tggctgttgg    480
cttcaccctt ggagcggctt ttgtctactt ggctgacctc ctgatgcctc acttgggtgc    540
agcagaagaa cccccagacg gccctggcac tgaacttcgg ctctacgttg atgaagaaga    600
agtctgatcc tgagggtccc gcgctactct tccctgagag tgaactttcc atccggatag    660
gtagagctgg gcttctttca gacaagagtg agaatggtga ggcatatcag agaaagaagg    720
cggcagccac tggccttcca gagggtcctg ctgtccctgt gccttctcga gggaatctgg    780
cacagcccgg cggcagcagc tggaggagga tcgcactgct catcttggcc atcactatac    840
acaacgttcc aaagggtctc gctgttggag                                     870
```

<210> 9
<211> 524
<212> DNA
<213> Homo sapiens

<400> 9

```
tgaaaatgca actagtctta gtggacaaaa caatcactga attaaaggag atcctgatta     60
tttctttcta ttcatacttt ttactttgtc tagaatgata ccatcaacac tccaaaaata    120
cttaattctg agcttttgca ttcaccccaa atttctcaca tgaaatgtta gcttatccct    180
gtttgtccct ttctcagatt gaatacacag ccgctgctca ttatctctta tgggatattc    240
ttgcttatat ttcgagactt tagctatatt cacagggtca gagagaggca tagtgtcttc    300
ctatctgtag gacagcagcg gcgtccctga attgacaagg agcatttttct tattgaattc    360
atctctcctt gccgaaagaa tgctatttag cttggcattt atgaaagcag atgttgaggc    420
cagaccacct acgttctatt tttgtgcggc ctcaggcaag ccaaccactc tttctgggct    480
tttttctctc aacaaaaata aaagtcgacg cggccgccaa atcc                     524
```

<210> 10
<211> 771
<212> DNA
<213> Homo sapiens

<400> 10
ggcacgagcg gcacgaggcg gcgggagggg aaacgtgcgc ggccgaaggg gaagcggagc 60
cggcgccggc tgcgcagagg agccgctctc gccgccgcca cctcggctgg gagcccacga 120
ggctgccgca tcctgccctc ggaacaatgg gactcggcgc gcgaggtgct tgggccgcgc 180
tgctcctggg gacgctgcag gtgctagcgc tgctgggggc cgcccatgaa agcgcagcca 240
tggcggcatc tgcaaacata gagaattctg ggcttccaca caactccagt gctaactcaa 300
cagagactct ccaacatggt gccttctgga ccatacaaat gaaacttcca acagtactgt 360
gaaaccacca acttcagtgg cctcagactc cagtaataca aacggtcact caccatgaaa 420
cctacagcgg catctaatac aacaacacca gggatggtct caacaaatat gactttctac 480
caccttaaag tctacaccca aaacaacaag tgtttcacag aacacatctc agatatcaac 540
atccacaatg accgtaaccc acaatagttc agtgacatct gctgcttcat cagtaacaag 600
tatgtattaa agctgacctt ttttttccta tgctgcaagt agttaaatgc cacatttaag 660
tcattattct gtctggcaaa aataatatag aaactttata tgcctagatt tcccttaatt 720
cctaattaaa tgttcttaac acagaaaatc agagagcttt aattacaatg t 771

<210> 11
<211> 1174
<212> DNA
<213> Homo sapiens

<400> 11
cggacgcgtg gggcagaagg tgcaagccag agctcaggca gaacttccag agtgcatctg 60
ggatctgcat ttgccactgg ttgtagagga ggcggtctat gatcactgaa aggcagatcc 120
atgtggctgc cccctgctct gttccttctc agactctgaa gctggttttt catccaaggc 180
ccagagtctg tgagagcccc tgagcaggtg accctgacgg ttcaatgcca ctataagcaa 240
ggatgggaga cctacattaa ttggtggagc cgaggggtgc gctgggatac atgccagatt 300
ctcattgaaa ccatagggtc tgagcaagga gagaagagtg accgtgtgtc catcaaggac 360
attcacaaag accgcacgtt cactgtgacc atggaggggc ttcaggcgag aagacccaga 420
tgtttactgg tgcgggattg aaagaagagg acctgacctt gggactcaag tgaaagtgat 480
tgttgaccca ggaaacaccc aggtctccct ggcaatcttc cccttaagct tcagacagac 540
ggctgctgat ggggccatct ggatggagac ttgggactcg ggcttgtgtt ttccaggctg 600
tttgtctctg acgggccccg gctctgtgac tggcactgcg ggggactctc tgacagtgtg 660
gtgtcagtat gagagcatgt acaagggata taacaagtac tggtgccgag gacagtacga 720
cacgtcatgt gagagcattg tggagaccaa gggagaagag aaggtggaga ggaatggccg 780
cgtgtccatc agagaccacc cggaggctct cgccttcact gtgaccatgc agaacctcaa 840
tgaagatgat gctggatctt actggtgcaa aattcagaca gtgtgggtcc tggattcatg 900
gtcacgcgat ccctcggacc tggttagggt gtatgtttcc ccagcaatta caaccccaag 960
gaggaccaca catccagcca cacctcccat cttcctggtg gtgaaccctg ggcgaaacct 1020
cagcaccggg gaggtgttga cccaaaattc agggttccgg ctcagcagcc ctcacttcct 1080
gctcgtggtc cttctgaagc tgcccctgct cctgagcatg ctgggtgctg ttttctgggt 1140
gaacaggcct cagtgggctc ctcctggaag atag 1174

<210> 12
<211> 1086
<212> DNA
<213> Homo sapiens

<400> 12
tatgtgatga tgcaacccca ccacacccac caaaagagga gtcttgtgga gcctaatcgt 60
tggcggcggc gacgaaaaat gtggggagga acagccacct tagtttcttt tctctgaaaa 120
gagccatatt ttatatatgg gccaaagctg aaaaaaacag tggcgcttat ttgtacaaac 180
gtaggacagt ggcgcatagt ggaaagaacg atagatctgg agttaaaata tctggatttc 240
tgtctgcttg catcattagc tatctgtgat tgtgagcaag tttcctagtg ttttaagtaa 300
ttcacccaag attatgcaag tagtgagttg cagaaccagg attcagatgc aaatctttct 360
aaccctaggg ctctatctct agtgtctata ctcctaacct agactgtggt gtttccccta 420
acttctgaat ctttctgata cagattttgt ccccatgaaa gggaacattc tcacacccac 480
ttggtttcat ttgttgctag taaagataaa atgggatcat acgtaggtaa agcccactgt 540

gcagtctaca ttatcatgat acttattgta taatatctga ttattcacta ttcaaaggct 600
ctacctctga gaactaggta taaatgaagt tagcttcttg agctctggaa atcaactaac 660
aaacacagcc cactcattgg agacattcat tcaacactat ttacttattt tttaaaagat 720
aaattgtatt gtgtatattt gaggtttata acatatgtta tgggatttat atataaatag 780
taaaatgttt agtgaagtaa attaacatgt ctgtcatctc acataggtac ttttcttggg 840
ggaaggcaaa aggcagctaa aatctgctta tttaacaaaa atacctacgt acattacaat 900
tttgttagcc atagtcgtca tgttatacat tagatctcta gacttgttca tcctacatat 960
gtgcaacttt ctatcctttg acctacatct ccccattccc tgccacctcc catttactcc 1020
taatttaggc taatttaatt ctgggttttc agcctataa catatcaaaa aatttttttg 1080
agtcag 1086

<210> 13
<211> 2169
<212> DNA
<213> Homo sapiens

<400> 13
tttcgtcgcc cggcgcggag agcgcgggag cggagcagta gcccgatccg gggccgcggg 60
ccgagctgcc gattggccaa cggctccttt caaccctgcc tcgttggtgg ccactggaga 120
agcgcggggg gctcccccag acagccgtgg ggacaagtta gagccagcac tttaccccgg 180
gccttgcgtg tagcttcccc tcccctactc tcggtgccct ggtgtctgga ggggggtcgt 240
gggggtgtgc ccgccttaca tggtccacca cccgggcaac cctctgggct tgtgttccat 300
ctcactcttg cttcctgtac tgtggtcaag gggaaccact gcatcatgtc ccggtatagc 360
taccagagtc tcctggactg gctctatggg ggcgtggacc ccagctctgc aggcaatggc 420
ggccccgact gtgctgcttt cctctcttgc cagcaccggc tgctggaaag tgtggtggtc 480
ctgaccctgg ctctgttgga gatcctggtg gccctgcggc acatcctgag gcagacgaag 540
gaggacggta ggggtagccc tggcagccag ccagagcagg tgacccagcg gccagaggaa 600
ggcaaggaga gcctgagcaa gaatctgctc ttagtagccc tgtgcctgac cttcggggtg 660
gaggtgggct ttaagttcgc caccaagacc gtcatctacc tgctcaaccc ctgtcacctg 720
gtcaccatga tgcatatctt tctcctggcc tgccctccat gtcggggagc tattcgtcgt 780
cttcaagtaa gagctacaga tgcacatgtt gaatggagct cttctggcat tgctgtttcc 840
tgtggtaaac actcggctgc tccccctttga attggagatt tactacattc agcatgttat 900
gctctacgtg gtacccatct acctgctttg gaaaggaggt gcttacactc cagagcccct 960
cagcagtttc cggtgggctc ttctctcaac tggcctcatg ttcttttatc acttcagcgt 1020
cttgcagatc ctcggcctgg tcaccgaagt gaatttgaac aacatgctgt gtccggccat 1080
ctcagaccca ttctacggcc cctggtatcg catctgggcc tcgggacacc agactctcat 1140
gaccatgacc cacgggaagc tggtcatcct gttctcatac atggctgggc ccttgtgtaa 1200
atatctgctg gatttgctcc ggcttccagc caagaaaata gactgaaggt gcttattttt 1260
ttttttttte ctccctgagg aagcaagtcg tgacttgact tggagaacac ccagttcttg 1320
ataaaatcat gggagagggc agtaggatgt ttggtgtatt tctttttcct cctttctgtc 1380
cctttcttct accactcttc ctttcccagc tcttcccccc acgatgtctc cttgagcctg 1440
gggtgtagtg ctgggcgctg gattttcctc cttccctttg gctctctctc tgctcctagt 1500
ggccttacat ggggagattg cagtgagcgg ctctttcact ctgcttgtcg gtgctttaac 1560
aacagctggt tgaggagaaa gggaacaaca agtagtagtt cttttggcat cgaagcaatg 1620
agattgggtt gtgtttcatt tctccactgc cagggacctc caggcttatg ctggcgtccc 1680
atttagtctc cagcacagcc cttaccacgg gggactgagg ccttgtttct ataccagcgg 1740
caagcccaga gccgagggat ttcacagctc gtggtcaaga cagaaccagc cagattctgg 1800
tcactggtgt cccatgactc aaagtaaaca gccaaccagc cactccctcc aggatgctgc 1860
ctggggtggg aggtgagagc accacccagc ccaaagggct agatttccat gagaaggttg 1920
gagatctatt gcaatgacaa atccaaagtg aagagggtca agatatatac agtactaggc 1980
cagaccctgt gtaagttctg cctgtgagag ttgagaactg atgcatatcc caatcagaag 2040
tgagcacaga aaactgtgat ttgaatcata cactctgttt taggaaaatg ttgataatag 2100
ataccatttt aatcttgagt gccccttaaa tgtgttccct ctccctttcc ctctcccccc 2160
acatgtatt 2169

<210> 14
<211> 2403
<212> DNA
<213> Homo sapiens

<400> 14
tttcgtagta cctgcagatc ctgacctctc tgatgagatg gatcgctttc gccgtcatga 60
ttgtgctggc cctgatccgc atcgggcacg gacaagggga ggggcacccg cccctggctg 120
acttctcggg ggtccggaac ctgtttgggg tgtgcgtcta ctccttcatg tgccagcact 180
ctctgccatc cctcattacc cccgtctcct tccaagcgcc acctcacaag gctggtgttc 240
ctgtactacg tgctgatcct ggccttctac ggcctcctct ccttcaccgc catcttctgc 300
ttccgcggcg acagcctcat ggacatgtac accctcaact tcgcgcgctg tgacgtcgtg 360
ggcctggccg ctgtgcgcct tttcctgggc ctcttccccg tcttcaccat cagcaccaac 420
ttccccatca ttgccgtgac cctgcgcaac aactggaaga cactcttcca ccgcgagggc 480
ggcacgtacc cgtgggtggt ggaccgcgtc gtgtttccca ccatcaccct ggtgccgcct 540
tgggctggtg gccttctgca cccgcgacct ggagtccgtg agtgggcatc gcaggggcta 600
tagagggcat caggcattca gtacgtcatc cccgccttcc tggtgtacct ataccgcagg 660
gacacccagc tggcctttgg ctgtggggtc agcaacaagc acaggtcccc tttgcgccac 720
accttatgag ggtgcttcgt gctgctctgg gctttctcct gcttcatttt cgtcacggcc 780
aatatcatcc tcagcgagac caagctctga tggcaggaca ggcaggtcta ctgacagaga 840
ggcatagggg ccccagcccc acccctcgcc cgcagagccc agatttagtt gctcccaggt 900
cttcatgggg caggtgggac tgtggctcta ggggaaaccc cctgtccagc tggggttgtt 960
ctccccaccc caccttccac ccagtctgca cctgtcctca ctccccgggg cagctctccc 1020
accaggtcct tctgcttagg tgacgtgtga actgctaaca agttctaaaa gtcattactg 1080
gaatcacacc tcttaaacat cttccaacac gcttcacccc caaccctact tccagggcgg 1140
gctactccct gcacaatggg agatgcagtg ctggcactgc catcatttgt cccagaccca 1200
tgcccctccc ctctcctgcc ccatcctgcg tccttgtcta agaaacctcc actagcagct 1260
gctgcctctt cagaggcgaa tcccaggctg gaagtatcct agccaggtct gcctgtgacg 1320
gcagggccca ccctcccacc aggcctggga ggcggtaggt gccacctcag agagggcctg 1380
gctggcccca tgacaagaac aagaagtgct aactggtacc agagtcccca gatagagctg 1440
cgagggccct cgctgtgggc gcgagctccg gagtagcagg ctgcccgtgc tcccactcca 1500
acactaaaag cggctccctt aggggctgag cgtgggtgct cagtgttcac gctatatccc 1560
ttgggcaatg tggggttgga tggggccccc acttccattc ccatggaagg aggccaggtc 1620
cccagccacc tcccactcag ccatgcacgc acttgctggg ctggcctcct gggaaacaca 1680
ggtgactcga atgaactctg cattttcaac gtgccttcta ctgcttcagg acctgggggt 1740
ccccctgacc ctcactggct tgccccccag ccctgggcct ggccccacct gtcctggagc 1800
ccagagcccc tggcctggag ctgcctcatc tgtgtgtgtg tggtactcag tgccccaccc 1860
ctccctcttt atgagttcag ttgccttgct cagcccctcc cctgtatctc agcgtctatc 1920
agacctctga cagagacgac gatgtcgtgt gtctccctgt gtggcccact ggtgtgtctc 1980
acgggtcatg gtgtgtggga tatgccggga actcggggag cccactggct gcgccacctc 2040
tgccctggca gctgaagcct gggagagtcc ctgcgtggtg taattggcct cagcccgctt 2100
tctctgtgcc gtcgcacctc agtgtttctc atagcttgtc cccacgtgtc actttccatc 2160
cacgggaaaa acaaatgccc cttctccatc tatcattgcg acttcctccc aaggaggcct 2220
ttcaggttgg gtaaagcagg ggcctccagt ggtcaggcca aaagcaagga aaacctggct 2280
gccccccctgt gggtgaaaac tcaaaacatt tggaactacc cttttccaat taggttccct 2340
tttggttaaa aacaaagcta attaaatcat cctgcccccc gagggtccaa accaaaaaaa 2400
aaa 2403

<210> 15
<211> 926
<212> DNA
<213> Homo sapiens

<400> 15
cggccgcgtc gaccacaagc aattattaga catcttctgt gtgcagggta tacagtagac 60
acactgagat ctggtcccat aatgcctcac cttaactgat gaaacagtat acagctagtc 120
tccctcctgc tgatgctctg tgcccccttt tccaaaaata atcctaaaac acatctcacc 180
acttcactct cttgatttaa aaaccaagtg gatcaattca ttcagtcaaa tatttatcca 240
ctggggatat aaaagtaaag gtgactaata ggccctcatt accctagact ttacattgtg 300
gggagtgacg ggggaagggg tggtggtggt gacaaataga agccaaggag tgtgcagaga 360
tttaaagcca ggcagacctt ggagaacacg acagcctgga tctcacaaat cagaagggca 420
caaacacaag ctggtcagag aaaggagccc tccaggcaca aaagccggcc cagaggccca 480
gggcagaagc tagcttggtt catgcaaggc aagaaaggag gctgcgctca tgagtgagaa 540
gccagcagtc tcagatgatg gggaggaagc aaggtaagta gttgggaatt tattccaagt 600

gtgacgggaa gccaatgggg aattctgagc atgggagtag cacagtccaa ctgccatttt 660
taaacattac tttggctgct gggaagaaaa tggattgcag gggagcaaga gagcagcaaa 720
gagccccact gggaggctgc tgccatttac tgacctcgca gtggagattc agccactcca 780
tgcccatctc caggttacaa gccctacctc ctttattaca gggcatcggc acaggccatg 840
tttccctgtc caccacccag ccacatacat aggcagccaa agccagtgcc atgcctctga 900
tctccaatgc ccacttccta gcgtcg 926

<210> 16
<211> 1460
<212> DNA
<213> Homo sapiens

<400> 16
ctgtggcgtc gtgttttgct tgctctgtga gcgtggtcca ggttcgacgc cttttgcgcc 60
ttagaggcat cagcgtatca gtttcctcgt aagtcgggag tgatggaacg accaccctgt 120
agtgcggctg tgagaattga gatgaggctt ggttcatatt gagtgctcag ctttcctggc 180
ttatagacag tggaggggcg tggccccacc atgcacccca agggtctcag actcaagggc 240
caacagtagg agttggcggg gagccttggg cccctctggc ctcaaccgga ttttccagcc 300
aaacgcagag agagatgccc tggaccatct tgctctttgc agctggctcc ttggggatcc 360
cagcaccatc catccggctg gtgcccccgt acccaagcag gcaaaaggac cccatccaca 420
tcgcatgcat ggcccctggg aacttcccgg gggcgaattt cacactgtat cgaggggggg 480
caggtgggtc cagccttctt ggcaggcccc cccacggacc agcgcggggt gacatttaaa 540
cctgagcggg ggcaacagca aggctttccg ggggaccctt ccactgccca gatggagagt 600
taggtgagct caaccagtcc cagctgtcag acactcagcg agcccgtgaa cagactcctt 660
accagggcgc actggatctt ggtgctctcc ctgagcctgg ctggagccct cttcctcctt 720
gctgggctgg tggctgttgc cctggtggtc aaaaaagtta aactcataaa tttacagaag 780
aaaagagtat cgagacatcc tgcgtggggc ccagattaac ttcgacagca cagacatgtc 840
ccttggaaaa ctccctgttt accggcttcc gcgaaaacga tgccagaaga agacccggcc 900
accttggatg atcactcagg caccactggc acccccaagc acttcaggaa ccggaagaag 960
gcccacttca cgtccttctt cgccgggaac ccccgaattc agcactttcc gggcctgcca 1020
gtgaggctga ggactgggggg acccctctgt ctccaggcat tcgggggcct gaggtccctc 1080
cagctacttc tgggggggct ctgtcagcca ctttctcagg gaattggaca gaggaaagga 1140
aggggaaccc tggccttggg attttcatca cagaggagtg ggagagggga cacaggcatg 1200
ggcctggcac tatacagaca accaggaagt tcccctctcg accttcggct cctcaggacc 1260
accagagaag gagatgtcag gaccctatct tgtcccccag atgggccata agacgtcaca 1320
ggtctctgca caacccgggg aattcctccc ttccccagtg ggtttttgag catagggtgc 1380
ccttgggtgt gttgtgtgtc tgcctgctgg cttgcttaag ttattaatta taacacgggt 1440
caaggtgttc ccaggcaaaa 1460

<210> 17
<211> 953
<212> DNA
<213> Homo sapiens

<400> 17
cggccgcgtc gacctttagg aaaaatatga aaaagtttca gctggtcttt gaccggaaag 60
tataggatgc agaagtcatc ttatagatag tgatgtgcct cattctcatt gattgtccaa 120
gaaactaagc agcaattgtc ctttatagca tgaagatttg tccatagctc tttaaacatg 180
tcttaggaag ctgaatgaaa ctcagcgttt gttagtagca cttggctcat tagttccttt 240
ccatctttgg caaataagaa attaattttt aattggggag gacttctgtc tcatcttgaa 300
ttgataactc ataatctaac tggggaacaa tccctggtgg ttagggagaa aacatacact 360
cgtttatctt tacattcata tattgcgtaa atgaccctaa agacgtaaaa ttcgcacact 420
aacatggttc tactgggaag aactgtggct atcaaacagc atcctaccac gcttgattcg 480
atgaatggaa taatataact gccatgattt ccattttacc ccccagtaga tctcaggact 540
cttcacctat tggtgacaca ataacagtac tccaggaagc tcagttaaac caccttaaag 600
gcagcaatcc tcagcaggcc cacggtagct gttaagctaa cagagcatga gagaccctgg 660
ccctcactac acatcatttc caggtatttc ctcccactgc gttagaatgc acggcctgac 720
aactgcctgg tgtgcaatat tggtaggtag tgggcgcaat gttagagaaa gaagtccttt 780

gagatggtag tgatactaaa ggcaggggag aaggactagg aacttcttcc agccttccca 840
agaaagaaga gccaaggggt gagttgttca ttcctttgct atctttcaag gtactatttt 900
aatagaggag tgggaaagtc acaactctga aagaagattt tcctactctt tag 953

<210> 18
<211> 1968
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(1968)
<223> n = a,t,c or g

<400> 18
caccttttca aaaaagattg ggattactat tttctgtgtc caggcttcaa cagtttaatc 60
tattatttac agctcctcaa gtgtcaccaa caaaaatata ctcttgtttt caatttctct 120
tggcactcag agatgaatct gtgtgagcac ggagagctag agatctcaga gcagccagga 180
actgagatgc aaatcatttc acctttgtgt ggtctgctca ttctgtagtc tttgtcttac 240
ccttaaagtg ggaaactctt tgacagggaa atgccatgtt tgaactgaag aagcctactt 300
tatagttttt tctctcttgc ctaatgtttt cagtagcgta ctatgttggg tgttttagta 360
tgccacatgg tctctattct ctgaagctgt atgatgaaac aagtacctaa ctttaacagg 420
agataattaa tgtcttaagg gaagtactaa cagaattgct ttagagtctg ggatttttat 480
atttttattg caagaactat ttttcatagg tggaaagcag aattcatatt aaagaaccag 540
tgggctcctg ggcccaccat ggcacattta gccttcactg ggaagtagaa ctaaatgaaa 600
gccatcagca gccaacacag cttcacttca gactcatagt atctgaatcc agtgttccca 660
aggacatcag tgactttata ttgttgcgca acctggtaat gcttaaaaac ttactttatt 720
ggatctcctt gcagcttttg acacagtgaa ccacattcct ttcctgaaat gctttcctct 780
cttggctttc tgatgccatg ttctcctgtt tcttctctac ttctctggcc acttctgtct 840
ccttagaagc tcagtcttgc tttgcctggc ccttgattgt tagttttccc caagggctca 900
cttcttagcc cctttctcct catgtcttat aattttgagt catctcatct actctgggga 960
gttgaatggt cgcttgtatg ctgaaaactc ccaaatttgt atctgtagcc cagcttttcc 1020
tactaagctc tacttacata tctttgcaga cctcattact tcttgactaa cttactggct 1080
ccttaaatat ctaaaattga atccatgttc tacctcccca tacaaaaagt tcctcctcct 1140
tccctaccac agttaacagt taccattact ctggccatta ctatcctggt ctctctgcct 1200
cctgatggga gtcagctcct cctgtcaagt tggcctccac agtggctttc acatgtctac 1260
cttccactct gctaatttct gagcagttac cacattcttg ctcaacagta ggtaccagtc 1320
acctggattg cttattaaat agttgaatta tggagcccca ggaatgtaga aaatttactn 1380
nnnnnnnnnn nnnnnnnnnn nnnntatggc cttagcaagt accacactgc actgtaagct 1440
ccattagcaa aggaagggtt atcctgcaag tgcagtatct aaggattatc attggttaac 1500
ttataggaaa ttgtcttcat ctgaacattt ctgatgaatc agagacatcg tacacaatta 1560
acagcaaata actgcccaat tacatatcta tcagatgagt gaaatgaatg actcctattc 1620
cacacacttc atgataacag ggctctgctg tttcaggctc aatttattta ccttgtcacc 1680
tggcccataa gtaagaagac tgtcatccat ggccattgtt ttctagttct tctgaggact 1740
tttctcctct ctttgtcagt ttaaagcctc cctaatcaac caaagttgtt tctaggcaca 1800
tgcactcttc cctcctctca tgggcctagc ctttggctaa gaagcccatc cttcgggggg 1860
ctcattacaa atttcacatt cttgcccggg cacggtgggt tcacacctgt aatccccaac 1920
actttgggga ggccgaggcg ggcatatcga ggtcaagaga tcaagacc 1968

<210> 19
<211> 1235
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(1235)
<223> n = a,t,c or g

<400> 19
aacttgttta catgtgtttg gcccatctgt ttaaacttct tgtttatttt aataggagca 60
actcttgggt gcaagctcct tttgtgcttg agaccactac tggcctcttc tccagctctg 120
tttccctaat ctgcattcta aatctcttct gtaaacaaaa cttgaataac aacttttttat 180
gataaatgct gaacacctcc aagatgactg ctgcaaaata aactactttt acaaccataa 240
aagcattcta aacacacaga aaatactttt ttaatgctgt ccttttcttn nnnnnnnnnn 300
nnnnnnnnnn nnnnncacac acaaagaatt ttatttttctg tgtgtgtgtg tctacatata 360
tatgttcttg acgatccttc tgaattgtgt gcattcccat catgctccag tgtttttcca 420
gccccgtaat gccctcactt ttgattgtgg agatgatgtt gaagtgcaca gggtatttac 480
tgatccggtt atcttgttct gtgacaggca gaatcaatca atcaagaggc cagctttgca 540
cttacatatt ccaggcacat ctccgcactt caggatacga gatcacagtg ttatccagcc 600
aaatttgatc aacaaatgca gttcccctgt gccacagcaa tctcagatga aagatgttat 660
tatgtggaac tgagcaaagt gataaagaca gtacagcaga gcctggagaa gcccggctgg 720
accgcactga cagcttttat atacgaaggc ggtgatgtcc ctttctttca gattataaaa 780
gttcccttga tttgaagaaa atcactctgg acatatacag gtgaccccca accattgcgt 840
ttcccatct ataactaaaa tcctcaagga agaagcaaca tacattagta taacattgca 900
ggcaaaatcc tttaataaca gggcagtatt taaagcgtaa agagagccac gtgtgaattt 960
tcagagagct tttttgatag ttttttgttg atgttgnnnn nnnnnnnnnn nnnnnnnnnn 1020
nggtgaaagc taatgttgga aagatgaaaa accatggttt tagaacttaa tcagttacaa 1080
ataattttat ttatatcaaa tctctgacac tgtgtgtcag tgtcaatatt ttaagtggat 1140
cccaagaaag tagtaagagg atgtcttact gcaccaatag caatcacttc attactgccc 1200
tgcagttatt atacattata tagtcgccgc ggccg 1235

<210> 20
<211> 333
<212> DNA
<213> Homo sapiens

<400> 20
tcccgttcgt acatttcgtt ggagggagcc gctctccaaa tgcctccacc ctgctgcctc 60
acttccactg agcacccctt tgacgaattt ccccaggtga caggcgtggg aattcttcgt 120
cccactgtta ctctttgatg cttcctcccc tgtgttggtg ttgtgttagg acgatgacat 180
gctgtattgg cacaagcacg gggatggatg gaagaccccc gtccccatgg aggaggatcc 240
cctgctggac acagacatgc tcatgtcgga attcagcgac accctcttct ccacactttc 300
ttcacaccag ccggtggcct ggcccaatgt att 333

<210> 21
<211> 1608
<212> DNA
<213> Homo sapiens

<400> 21
gagagagcgc gagacggggg gagcggaggg ggaggaaggg ggggaggggt aagggaagtg 60
gggaggaggg ggcagaggtg agaggaaggg cggggagagg gggagaagga agaggcgaga 120
gaaagggata gaagagagga cgagagaggg gagagaagca gtaacgctaa gaccatgact 180
cctagagtta ttaattgtct ctttcaagcc accatgtatg cagaattcaa tttggggcaa 240
agaagatggc aatcaaggaa tgcacccaac ctttctggct ggttggggct ggcaggggct 300
gctccctggc aggggagaat cagccccatg ctagggacca aggtatccct gtgtaacttg 360
tctgaagagt ccttagctcc attggccaag catactccca gggcctagcc tggggcatgg 420
aacactgtgg gtgcctaatc aatgtttgtt gtgtgaatga atgactaaaa ggagagcaga 480
agtgttgcct cagttcggga gctgggggtg cacctctccc ccagggctcc tgggtggtag 540
tcacttaaat cttgtctggc cccgtcctgc ccagcctccc ctaggcttca gcagggctga 600
atggcagcgt aacgcgtgaa caggcggaga gctggcgcta gcaattgaaa catctccggt 660
cagccggggc ccattgcggg ctcttcctgc ctctgctgcc cgctagcttt gctactcttt 720
gtgttgctgc ctcccaggtg gccagcgtca tgcccatagt gtgtctttct ggcctgctct 780
ggtgtctcag aagtggcctg gaaatgttct gctgattatt acgctctcca gtcaccttag 840
agccaggaag gcagttgcag aaagtcaagt ggctttggtc tcatctgctg gacttgttgt 900

agttaactct tggttagtcc tcattttctg agcacgtcct cagaacccac atcagtgata 960
gcagtgctgg ggtggtagca gtggaaggga cacgttgcca cctctggggc tcatagccct 1020
ggtggaatgc agacgagtaa ccgtggaggg gtagaagtgt gacatggtgg tcaggacgta 1080
cgctctggag gcaggcagtc ctcacttcaa agctcggctt tgccacttac tcaatgcatg 1140
ccctattcat gttgttctcc ctccctgcgc ctcagtctct atatctgcag aatggggaca 1200
ataccccttc tgcttgcctc atacagcgaa gggggtcttg cttaatggtt gagttttgga 1260
tctcagatca ggtgagagag ggctctggga aagtgacgga ggggtcagac ctgggccagg 1320
aagctatctc tggggcccac ccagcccgtg attggggatc cctcctgcat ggtccatgta 1380
ttcctcgcgg gaagcgaggt ccactctgga acccaatggc tccttgtagc atttcaaaat 1440
gccaaggcag ttggaatgtt taaaggaaaa cccagatgcg tgctgccctt ctcctccctc 1500
ctagcctcca gttataatgt tctagaaggg gcagaatttc tctcgtgggt cagccccttg 1560
cagcctgttt tatacagctg cagtcgacgc ggccgcggaa caacgaca 1608

<210> 22
<211> 1245
<212> DNA
<213> Homo sapiens

<400> 22
ctcacaagag agcagaggag ccccaagtct tggggaccac agaagatgcc atgtgctcca 60
cgatgtcggc ccccacctgc ctggcccact tgcctccctg cttcctgctg ctggcacgtg 120
gtcgcttgtc tccctacagc atgccgtctg ggcagcacgc agctaggtta tgatctggca 180
gtgtgcttac atgccctgag ggccccccatg ctgtcggtag aagtgtgaga cacttctact 240
gaggtgtatg gaggccggct cctgcactga tggtatgata aaatgggtga aggtgagcac 300
tcaggaccaa caggaaattt ataactttaa acgtggctcc ttccctgggg taatgcccat 360
gatccaacgg acatcagaac cactgaattg tgattattcc atctatatcc acaatgtcac 420
cagggagcac actggaacct accactgtgt gaggtttgat ggtttgagtg agcgactcag 480
aaatgaaatc ggatgaaggc acctcagtcc cttgtgaagg gagctgggga ccctgaacca 540
gacctgtgga tcatccagcc ccaggaattg gtgttgggga ccactggaga cactgtcttt 600
ctgaactgca cagtgcttgg agacggtccc cctggaccca tcaggtggtt ccagggagct 660
ggtctgagcc gggaggccat ttacaacttt ggaggcatct cccaccccaa ggcgacagcg 720
gtgcaggcct ccaacaatga cttcagcatt cttctgcaaa acgtctccag tgaggatgca 780
ggcacctatt actgtgtaaa gtttcagagg aaacccaaca ggcaatacct gtctggacag 840
ggcaccagcc tgaaagtgaa aggtgagaca attggttttg gacttgtcgg gagagcggag 900
caggtattca acaggaattt ttctacatgt cccgaacaag ccagttcttc acaggcctcc 960
atgtctttac ataagttttt ccctctacct ggaatgcttt ctccttttct ctgaaaactc 1020
ctactcatcc ttcaaggccc agcacaaatg gcccttcctc tgtaaaacct tctcctcctc 1080
taaagcagga gtaatttctc cccatctggg tttcaacaga actctgtata gatcattaat 1140
aattcaacaa aacattggcc agggcaccgg tagttaactc ttgtgaattc caacaccttt 1200
gggaggccca aggcaaggtg gatcaccttg aatcaaggga gtcaa 1245

<210> 23
<211> 374
<212> DNA
<213> Homo sapiens

<400> 23
aatacaccag ttgactcatt cagtattata gctctgagta cctactacat accaatggta 60
ggattgggga acatgaaact caggaagaca caagtgcccg gccctgaagc cattctctgt 120
cttccagtgg agatggtcaa aaaattgccc agaggcggaa acccagaaga tggcaggact 180
acagaggagg caccatcata gctcggaaga agccagtgaa gccttcctga atctggcatc 240
aagctaaaat gaactcccac gtacagacac tctccatgtc accagtttat ggggtgtttc 300
cacaagcatc cattctattt aatcacccga atgacccttg aagtaccaat tgttttggct 360
gtttttaag cacg 374

<210> 24

<211> 569
<212> DNA
<213> Homo sapiens

<400> 24
cttacaggct taattactta cttagttaaa tatttctgaa gactattcca ttcacagtct 60
ttttcctgaa taagccactg gatgggtact tacttgacct cctaggcagt ttgaaagtgg 120
gaaataggaa ctcaggttga tgagcattat tatttctgtc attagacagg tatacttctg 180
tcttatttac caatcttgtg taattgcatt ggcttactaa taaaaactta agtctaaaca 240
atgttcttct ggcgtctact catactgtat tctttctcctg agatcactgt ttgtctacca 300
tcttttcaca tctagtgggc tgaaaatgca gcatttgaga agtcactgcc atctctttcg 360
tcaggctctc tttctttgct cacacctaca cctcatctcc tggctatctt cttatcttac 420
tggtgtcaat tttctgtatg aactcctctt ctaggaccca tccttgaaat gttggagagc 480
ctcaacatgt ggtcccaggc tggcgtttaa gaaaacttcc attcctcctg ttccagtttt 540
taccatctca gtaagggtcg acgcggccg 569

<210> 25
<211> 842
<212> DNA
<213> Homo sapiens

<400> 25
gcgcaacagt gtaggccata gaggattact aggtctttac aggcttattg aagagtaagg 60
ctgagaatat gctggatgaa aaatgagccc aaatccacca tctcctgtcc ctccagttat 120
gtgctctggg gcaatgtggt ccgtcacaag taccatcttc atatgcaggc tactcattgt 180
aagactcttg ggcaacacag cagtgagaac ctctgttgtc tttctacccc acaaggctcg 240
cactgcattg tcataagacc actaccctgg ggtcaggagg ctgaggttca ttcctgctct 300
gtcatgaatt gcctacatga ctgaggacca aggacttgag ttttctcatc tatgaaacaa 360
agatattgga taccaggcat tttcaaacat ttctttttcc aataagcaga acccttttca 420
aatgaaatct tatgtatagt gcagttgctc agagccagga gctctcggct gccctttcac 480
cctttatggt agaaatcttg aggagcctcc acaagcatca cagcactggt cagtacttct 540
gcacatgagg tctgaggacc acctgcatct gaatcaccca ggtgcagatt cctgggctct 600
ttcccggacc tcttgaatca aacttttttgg gtctgggact gaggactcta tttttaatga 660
gcacttcatg ggattctcat gtacattgaa gtttgagtac cactaggcca gatatttcaa 720
ggttcctctc aggctttaag tacagtgact ctgtgatcct cttcagcttt ggtgtgatgg 780
tgaggagcca aagaagggga gtcagtgact ccacctgata aaaagattta acagcatgta 840
tt 842

<210> 26
<211> 915
<212> DNA
<213> Homo sapiens

<400> 26
cggaattgaa cagcatcttg gtgaggacat tttggttgag cccagaatct atcatgcatt 60
aacgataata gcgtcgccct gtacttacta taggaatgta ttatttaaaa aagtgttagt 120
aatgtcaaag agccattttt caaatgcctc tttttcatta ggaaaaaatg ggtggaaaaa 180
aacctttggt ttttcccccc aatcaaaaga taagggtttc tttgtccatg aaatgaagga 240
gttgctgaga cttctaaaag ctgccccacc cagggaaaga aagaggaagg atgctctctg 300
tgctcttgcc ctctttttttc tttaaggaag acaacagtca agacagctcc acaagcccca 360
gccctaactt taagtacctt tttgtttttt tgaagatagc agaaaagatg accgccatat 420
ctttaacttt taccccaagc ctccttttat ccttatggtt tccggctttc catgataact 480
acttttccat taaagacaca tgaagtcttc ttgaatgttt gacaaaataa atgtgttggc 540
tttcttttat cccatcaaaa aggcattttc tcctatcaaa atgcattttc atgtggaaca 600
tctgtgcaca gccaggttaa atttcaccac ctttctttta cctaggggtg gccatgttgt 660
catgaatgga atgtaatcag caaacttctg tgcctctgtg gaccaaggct tgcaggtggt 720
agggggtggc acccccatgt tcttattagc aagatgtggg tggtgccaca gctgtagaga 780

ctggtggggc cacaagagga gagatgccag gatactgtct gcccacatgg aggaagacag 840
caccctgccg ggaattgtgc cttaaagtgc tccatgggca aaaaagcaag tatgcatgtt 900
ttgtcgacgc ggccg 915

<210> 27
<211> 863
<212> DNA
<213> Homo sapiens

<400> 27
cgacctccat cctagctata aaggtcttat gagacttttg cactgtaaaa ctttacacat 60
tgtgctattc actctgcttt acaagatttt gatggatcat caaaatctgt cagaacatgt 120
actctgcatg gttttatatc tgattgaatt aggacttgaa aattctgctg aagaagaatc 180
agatgaagag gcatcagtgg gtggaccaga acgttgtcat gacagttggt ttcctggcag 240
taacttagtg tcaaacatgc gacactttat aaactatgtt agagtaagag ttccagagac 300
tgctcctgaa gtaaagagag actcacctgc aagtactagc tctgataact tgggttcttt 360
acaaaattct ggtacagctc aagttttcag tttagtagca gaacgtagaa agaaatttca 420
ggaaatcatc aatcgcagta gcagtgaagc aaatcaggtg gttcgtccca caacttcaag 480
taaatggtct gctcctgggt cagctccaca gttaactaca gccatttttg gaaattaaag 540
aaagcatatt gtctttgcta attaaacttc accacaaact ctcaggaaaa caaaactcct 600
actatcctcc ttggcttgat gacatagaaa ttttaatcca accagaaatt cctaaataca 660
gtcatggaga tggtataact gccgtggaaa gaattttact aaaagctgca tcgcaaagta 720
gaatgaacaa acgcatcatt gaagagatat gtagaaaagt gacccctcct gtaccaccta 780
aaaaagtcac tgcagcagag aagaaaacat tggacaaaga agaaagggtg ttaatgggag 840
tagtaggaca taaggctaag aag 863

<210> 28
<211> 980
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(980)
<223> n = a,t,c or g

<400> 28
cgacccacct ttctcctttt ttccgatgac ttccatttct atcaatggca gaaccatttt 60
ctttgtttct aagtcttgga gtcacagtca cacacatttt cagtgatgaa aatgtttcag 120
atgttactga cttcaagctt ttgctcatta tcccatctcc aatcatgcca gcacatttca 180
ttcctgagca tatcgaacca tagtaagata tttctttact tgcagcctac ctgttatctg 240
tatctacccc cactcccact cttttcacga tcctggcatt ggaacctcag agtccacatt 300
tgctcccccc gaacggaggt atgttcccca aggctgggcc tctcctgttc tttcaggcag 360
tcttcaggtc aatgtccttc aaaacgtggt gctcagaact aaacacaaaa ctccaggacc 420
actggagtgg gacagtggat ggactgcact tccaacaaaa cactgtagat ctccgaataa 480
acgtaagaat ggccgctata agcacgcacc cggcctaaag ccctggtgaa agctcaccag 540
cagtggggct cagtagctct gcgaggcggg aattagttct cttacttgac atattaggaa 600
atgtgaaatg tagacagatt aaagcnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 660
nntagaattc tgcctaaccc agagtctgat ttcagagcca ctgttctttt taccacattg 720
tactacatgt gaaattgtat caggccgggt tttcctatcc ctcatcttat ttacatcaaa 780
gagtcagagg ggatggggac cacaatccca ttagggaggt gcttctgaat cactcccagg 840
ggacactttc aagctccacc ctcctgctat acccagtcac tgcttacaca gactctgcat 900
tctccctggc ccactgtgct ccagccacac ttccccgtcc ttaaacagtc gacgcggccg 960
cgaatctcag aacaacgaca 980

<210> 29

<211> 2199
<212> DNA
<213> Homo sapiens

<400> 29
atggaggagg agctgcagca ttcccattgt gtgaattgtg tcagtagacg gtgcatgacc 60
aggccagagc cagggatttc ctgtgatttg attggttgtc cattggtttg tggtgcagtt 120
ttccattctt gtaaagctga tgagcatcga cttttatgtc catttgaacg agtgccttgc 180
ttaaatagtg actttggatg tccatttacc atggcccgaa ataaagttgc tgaacatcta 240
gaaatgtgtc ctgcaagtgt ggtgtgctgt actatggaat ggaatcgatg gccagttagt 300
tatgcagacc ggaaatcata tgaaaatcta agcagagatg tcgatgaagt ggcacaattg 360
gatatggcct tggctcttca agaccaaagg atgctcttta gaatccctca aagtagccac 420
catgatgtca aaagcaactg ataaagtatc caaacctaga gaacaaatct cagttaaatc 480
aagtgtccca gaaataccac atgctaatgg tttagtgtct gttgatgaag aatcttatgg 540
tgcactttat caagctactg tagaaacaac cagaagtttg gctgctgctt tggatatcct 600
gaatactgct acaagagaca ttggcatgtt aaatacaagt gtcccaaatg acatggatga 660
acagcaaaat gcgagagaaa gcttagagga tcaaaacttg aaagaccaag atcatcttta 720
tgaggaggaa ataggagcag taggtggaat tgactacaat gacacaaatc agaatgccca 780
gtctgaacaa aatggttcaa gtgatttatt atgtgacttg aatacaagtt cttatgacac 840
ttctgctctt tgtaatggct ttcctttgga aaatatatgt acccaggtca tagaccagaa 900
tcagaattta catggtgatt caaaacaaag taacttaaca aatggagact gtgtggcatc 960
atcagatggc acttcaaaac cttccagctc acttgcggtg gcagcacaac ttagggaaat 1020
aataccatcc agtgctttgc ctaatggcac agttcagcat atcctcatgc cagatgatga 1080
aggtgaaggt gaattgtgtt ggaaaaaagt agacttaggg gacgtgaaga atgtggatgt 1140
cttatctttc agtcatgctc cttcattcaa ttttctttct aattcatgtt ggtctaaacc 1200
aaaggaagat aaagcagtag atacatcaga tttggaagtt gcagaagatc ctatgggcct 1260
ccaaggaata gatctgatca cagcagcatt gctttttgt ctaggagatt ctccaggagg 1320
gaggggtata tctgatagcc gcatggctga tatttatcac attgacgttg ggactcagac 1380
tttttcactt ccatctgcaa tattagctac aagtacaatg gttggggaga tagcttcagc 1440
ttcagcttgt gatcatgcca atccacagct ttcaaatcca agtccgtttc agacacttgg 1500
gctggattta gtattggaat gtgtcgctag gtaccaaccc aagcaagcgt tcaatgttta 1560
cctttgtgtg tggacagtta tttagaagga aagaattttc ttcccacttt aagaatgtgc 1620
atggtgacat tcatgctgga ctcaatggct ggatggaaca gaggtgccct ttagcttact 1680
atggttgtac ctattctcag cgtagatttt gtccatcaat acaaggagca tagatcatac 1740
atgaccgcca tttgaggtca tttggagttc agccatgtgt atctacagta ttagtggagc 1800
ctgctagaaa ctgtgtgttg ggattacata atgaccatct aaatagtctt ccttttgagg 1860
tcctgcagca tattgcaggc tttctcgatg gcttcaacct atggcagctc tcatgtggat 1920
cccagttaat gagggatgtg tgtggcagcc tggcttcagt ctcgtggcat ggtcatgctg 1980
cagtggggta aaagcaagta tccagaaggg ccattcatca tggcagataa aacaaaagtg 2040
tatggcgatt ttagtgtctg cagtcttgtt gttggtaatg aatggtcatc tagagccatc 2100
gtaagcatgg ctgtgcgctt gaagcaacgc agctaccatg ttgtcaagaa acgggagagt 2160
gcactacgtc tagcttgtat gtgtgtatca ggtcaactc 2199

<210> 30
<211> 648
<212> DNA
<213> Homo sapiens

<400> 30
tcgttgttct gaggaattcg cggccgcgtg cacttttgct ttgtttgcag agtacgttag 60
tatatatcca agtactttat agtagattga taattctgtg aaggaaatga aggcgtgagt 120
atcagggtaa cccacatgaa agtttcctga ctaaaggtag ctggtttaaa cctacaaaca 180
tttacatgtg attggtatga tgggtactga atgaattggg ttggatcagt ataagttgat 240
catgtcatgt agaaagttat taaagtaggg gaaagaggtt tgtactccag actggtttag 300
ggcaggtttt tttgaattaa aaacaagtgt ccaaggaagt tgcccagatg agcagtacta 360
ggtcatgctg gtgggctgaa gctggagttt ggacctgacc ctcaaacaca ggtgcatgtg 420
cacatttcgt ggtctcctta ctgggctttt gacttttccc ttatttttctc ctgtgcttta 480
tttctgtaac aaatttccca ataaaaccaa catgttcctt ttatgttttt gtaaaaatta 540
tttcttaagt actgtttata aagtttttgt gagacaaggt ttggtttttg ctgcccacac 600
tgcagaggca atggtcgcag tagaagttgc ccgcaaacgc acgcgcgc 648

<210> 31
<211> 358
<212> DNA
<213> Homo sapiens

<400> 31
gaaatgcctc tactgcatgg tcacctttt ctccaagaaa gaaaagattt ctctctagct 60
ttaataagca gagctttgca tactaaataa cttggatttt cccatcatcc caacattcaa 120
accaagcctc agaggaggaa aaatgcttca gagaggctgc aatgactggg ttttgccatg 180
tgtttctgag aatcagaaag gcactgcaga tcttctagtc ttctggagtc ctgtggtctc 240
tgctgagcac tgccactggt gccctgattc aaggcccaag cttgggaaac attgtttctg 300
gagctgctaa accacaaccc agaaaggagc acaagggagt gatgaagggc atcctctt 358

<210> 32
<211> 474
<212> DNA
<213> Homo sapiens

<400> 32
cttacacggc cgtgctgact cagtaaatga tggcggcttc agcaggaagc ccagggtgtg 60
ctatggcatt ggtggggtcc ccaaccagat agcctccagc gccacagcct tttacctgca 120
gcttttcctg cttgatatag cacaggtaag tgtgggcagt agcccagtct ctgctggcat 180
cttccttggg gtctggtcct tccaccccca cctctgcttc tgctcccact atccttgtag 240
acagcacttc tgacacgagc tcaacaaaaa gccagaagcc aatgagggca aggtactcaa 300
gaatgtttac tgcagcttct tcctgtggtt cctgcccccc ttcaccagcc tgcgaggcct 360
ctggtacacg actttctact gcctgttcca ggccctggcc acggtaagca gggccccttc 420
ctgggcctgt gctctggcga agcccattgc tggccatggc cactctgaag tgtg 474

<210> 33
<211> 1275
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(1275)
<223> n = a,t,c or g

<400> 33
ttatatatgt aacataatga agccaactcc agggctctac attaactctt tccctgtagc 60
cccagataag gaaaaaaaga gaattgttca agaatacatc tttcctacat acaaatcagc 120
aatcaagaca gatggagagc tgaaatctag catcctgaat ggcagagcac attgactgca 180
aaaaaaatac ttaaaacaat tcaacaaagc agacaaaaaa aacctgctta aatatatgat 240
taagaagaac atattgatgc caaaaaccaa caagataatc tcagtggatg acctcctttt 300
gaggctacaa ggctagtcca tatcaaagca ttcattcatc ctgtagaccg ataggtatgg 360
aggggaagta ttattggaga gacctcaggg ttttctcacc tgagaaggga gagagccctg 420
gggtggtcag agctgggccc tgagtggtgg gatttagaaa gaagggaaga ggggagggct 480
tggaagatag gagaaacagc atgaacagca gcacggggaa tgggtgagtg tcctcctgaa 540
aggactgtgg agatggccgt ggagcctctt cttgcccact tcctgagatg gtcatggctt 600
agtgctagag acttttattc cttgggcaat gtagacccag ctctatgggt tccctgtttt 660
ttccttctct tcctcctcat tataacagat aataataatg attcctagct ttcagggaac 720
acatcttgta taccatgttc tgtgtgcttt acatggattg tctcattaca gcttcacaac 780
aaccttagag atagatgttc tgtatcatta tgtgcacctt acagttggtg cagctgaggc 840
gtagagaggt taaataactt gcccttggtg ggtggcatag catagacagt agttgggtga 900

ggagtctacc aaggcagcct gaaaggcagg aactcagtga cagcatcagg gcatgctgct 960
aacccactgc ttcccactcg tctgaatatt agaatcacct tggaatcttt cagaatttcc 1020
aaatcccagg tcgattacat cagagtccct tggggtgctg ggacacaggc atcagttttt 1080
gaggcttccc aggtgatccc tgcccaggnn nnnnnnnann ntnnnnnnnn nnnnnnnnnn 1140
ncntnnnngn nnnttgnngt ggtggngtac gaggcgtgag ttgatgggtg gttggtgggg 1200
tttgctggtg gtacgattag ggaaggttcg gcaactggct gcgatccctc gccatttcgc 1260
cgatacgacg tgccg 1275

<210> 34
<211> 862
<212> DNA
<213> Homo sapiens

<400> 34
cgacttacag tctgtctctg aagattactt gtacagtccc atcagctata ctttttcacc 60
ccacagtcca gctgccagcc tcgccagaat ccccttatta tttgaaattt attgttgtat 120
tcactcaatg tattctgtgc ctgaaatgca tctctggagt caacatgggt acccctttcct 180
tctttaaagc ccttccctga tgtgcagtgt tacctgtggt gtcctcttcg ctctctcagg 240
cctcctgtta tactcttctc cttccccaca ctggaacaga cccagtagaa ttgcagtgta 300
tctcatgcgc ctcactaagt actgcactgg gagctctgcc gcctcctgcc agtgactctg 360
cttcagccat gttttcatgc acagtgactt gtcacatttc tcacacatgg cagacattta 420
taatgtctga tgtgtgttga gtgtctcaaa agaaaaaggg acacaggtag accttcagca 480
aggttgatat ctagttagtt gacagttttt tgttaagtat tgcttcagcg agctctacat 540
aaggggatac agtattttct tcatttatta ttttctctcc ttcagggcac tgacaacctt 600
aaataggtca ctttaaaaaa aatagctatt tcctatttac ataataaata catgtgaaga 660
tttaagtttt gtaagttgag gttgaaaagc tgatttaatc actgtggcag gagttgcaga 720
taataattag gtgaaagaat aggaaggtcc tcattgataa cccattaaaa tgagtattta 780
ggactgaacc aagagaggtg attcatcttt ctgtctgtat agaatttgga taattgactc 840
agtaatgaaa tggaaccaat ca 862

<210> 35
<211> 761
<212> DNA
<213> Homo sapiens

<400> 35
ctggttctct ctgttgcaca ctgagcaaca ttgtcttttt aaaatctatt atgaaataaa 60
gttaaacagc ttcataatta tcattgtcat tatttttctt tttagctact atttttggt 120
ctgtcctgtg gaaacagctt gcaatcctca caagaacctc gtaatgtgaa tatgccagcg 180
tgttgctaca gaccgtgctt acttcagcca atttctctgc tgaatatcct ccttttactg 240
atgaggaaac caagtcaaga agtcataaat gacacaccca aggctgggaa gtggttgagc 300
aggtatttag actcaggtct attttactcc tgtgcctgtg ggccgttttc acattctttt 360
cagttggaca gttttaaaaa tgtgtaatgt tgcactcaaa ttaaataacc tgaaattttg 420
caattctgaa agggaactga taacttgagg cactatgtaa ctatgcttca acacctcaaa 480
aaggctccaa cgagcacttc tgaaaatgat ttttaaagcc ctttttttact tgaatgctgt 540
tcaaccttct caccacttgg gctaatgttt tatgctgaac caaagaatta tggacccttt 600
attctggggg ctggattact tgaatccaaa gtgatgtgat gaataacaag gatctaggcg 660
ttgttgccca gttgatagtc acaaaatact tggccttaat aagtcaggaa ttattttccc 720
catatttctc gaattacatt aaagaatata aggtttttaa a 761

<210> 36
<211> 614
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> (1)...(614)
<223> n = a,t,c or g

<400> 36
gcaacctggg tggatttaag gtggaggtct tcctctctcc aggtggctgc cctcatctct 60
acccccaccc caatcccatg cattctcctg ccctcctgct ttgggcaaga cccgggcgat 120
gtggagttgg catgtccagt tgcaggtatc agcaccactg caccatctcc tctgccttca 180
cttcccacca gcacacagaa tttatatgcc cttcccaagt ccaaaaaggg cccctgccnn 240
nnnnnnnnnn nnnnnnnnnn ngcaaggcat gagctctgtc tcatggaagg gagaggccaa 300
gttctctttc caccatcaga gagtagcgtt caacatcatt tatacaagac aggcttttgc 360
ccttttagtc ttgctgaact agaatgaggg atttgtgtag tcgaaagtca agctgatcat 420
tgtcaagttt cgtggaagtg gaacctgaag gaaggttagg tggagactgc agagagggtt 480
agcgtcctgc cttaccacaa ggtaagtcta gaagagggtg gggaagggag ttcagaattc 540
ttcattcatg gagtgtacct gccttggtca tccttttatc ttccagtcag aggggtcatc 600
agtcaagagc tggc 614

<210> 37
<211> 387
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(387)
<223> n = a,t,c or g

<400> 37
gatgattgct gtagggtttc accacgggca actctaactt gtcacagtca atccacacct 60
tgtcagcatt taaccgcctc acnnnnnnnn nnnnnnnnnn nnncgaggcc acagctgttc 120
ctgccctctg tgttctttgt gctgttattc tcatatactt tcacagaaac cacacaatgg 180
actgttgtga ttttggcttt aaacagtaag ttatctccta aagacattga cccaatctgt 240
tagtctttta gtcttttatc cttccctcta ccatttctga tgtgcttcat tcctttgtgt 300
ggacctgagt ttccacccgg catcatcttt ctcctgtgtg aataacttcc tttagcagtt 360
gttatagaac agatgtcgac gcggccg 387

<210> 38
<211> 850
<212> DNA
<213> Homo sapiens

<400> 38
cccggtcgac gatttcgttt tgcctttaca atagttatca gctaaacctt tgaattaaaa 60
aaacaaacag tatctaagct ttattgagca atgtttatgc caggaactgt gctaaggatt 120
ttactagcat taccttattt aatccttaca aagcaggtac aatttttttct attttctgat 180
gaaattatgg catggaaagt ggtggcaccg gggcttgaac tcagtgctgt gactccagac 240
tccacactct ttaaccatta cactatacta tcctgacatt aatattctat atgattaagc 300
ttgcggagaa gtaaattgag gtttgatatg actttacacc gttttaatca atacctggtg 360
ggttatggaa aggcacagtg ttatcaaggt gtataatatt ggggagggta tggatttcat 420
ctagttacta cacaatctta cagtgtcaga gctgatccta gtcctccttg attgtctttg 480
gtgctgaact tactgatcct tagccctaag tcagtgtagc agagtggtta aggacagggg 540
tcctggggca agaccagggt ttgaatcttt atcttagtca aaatatctga cttctatgcc 600
tcagtttttt cattatactt ctcgaaaaat gcatagccct tcatctctct ctaaaataac 660
atcccttttgc atttttattt ttgtctcttt ctgagactgt tactatctga atgcctgctc 720
ttctagcctc catatctttt agcttttctt ttgcattcct atttctttat tctttcttac 780
tatcttctgg aagagtatct cagtcttatc tcctaattca ctaagatatt caactgtacc 840
ccgctgtatt 850

<210> 39
<211> 612
<212> DNA
<213> Homo sapiens

<400> 39
aattcccggg tcgactacga tgccgacgat ctcgcgcaat tctcggacgc gtagacatga 60
ttaaagagat tcatctggtt ccttctgtgg ccggcgactg catttcgtca gaactgaata 120
tgccatgata ccttgcaaat taatctctat agcacaggac agcttcttaa aaaggctcat 180
tcagtttagc aaaattgaaa acccagtttt agttctgaaa attaatcttt gcaaaacaac 240
aggaaaagct ccccacccca cacctggcct gtaacaatct tttcatctag cactattgta 300
aagaagggtt tgagaatgga aatgggtaag tgcagtccta cttcactttt ctttctcatt 360
tccatagtag tgaaaatgta cagaagagac atgtttactg tttctccatt ggctgttaca 420
gttttccaga ggaatgaaat gtaatattta agaatgaact tgacttgcat gaattaaaat 480
gttttcaaaa ttttgtgtgc taagatgacg gcagccaaaa cagctatgtt tgggttatct 540
tgtttttaa attgcatgtt ggacatccaa agaatttcct cttttatctt gcgtaggtat 600
ttgttgactg ac 612

<210> 40
<211> 600
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(600)
<223> n = a,t,c or g

<400> 40
cccgttctga cgatttcgtc ctagcgccca tctttatgtg gggctcccta gaaacccttc 60
tcttttattt gccacttcct gttttgtgga cactccatgc ctgagctgcc aacgtgggtg 120
ctggcattac tgcctcatcc tgtggttttg ctcattgact ctggagaact agaggccttt 180
gagcagatat gcagaagcac cttgaaggca gtgtggcact cagtacatgg tgccatgtct 240
gtgtgcattc atctgcttca ccttttgcca ttaaagcagt ttccagggga ctgtattcca 300
gcatagtgtt gttctaaaat cttcttactt aaggtgattt ttacaacctc tttgctttcc 360
ttcattgccc agcaagtctt tactcttgtg aaatatctta ttgcagtcta gcccaaatgt 420
tgtactttta tttattcttg tttcttttaa catgccatag aagagtttaa agctgtctgt 480
taggagatag aattgatnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 540
nnnnnnnnnn caacgatctt ctgacgttaa aaccatgaga cccaccttat caacctgacg 600

<210> 41
<211> 348
<212> DNA
<213> Homo sapiens

<400> 41
cgcgttagcc ctgaattccc tgacctttcc cttttgggct cccatgagat gtgttaggag 60
aatgtgcatg tggagtgcgg gcccagcccc tgctaccgtt tgcgctgtaa tggtggctgg 120
tcctaagagt cctcagtttc ctccccgttg gggttgtgtg tacaacctta tcggttgtca 180
ttttttggat ccattttcag gttatttttc tggtattagc tggagggaca tttcctttag 240
cctgtacttt atgggtcagg agtttcaaaa ccagtcccct ttgaagtgag agtatcccta 300
ataaaaaggt gagtgtcccc cctcctgtgc cctggttttt ttggttgg 348

<210> 42
<211> 379
<212> DNA
<213> Homo sapiens

<400> 42
aaagaagaaa gtgaatcctc ctccattgtt tcaaatataa aactggtatt tcccttatag 60
aatactgggt tatttttaga atacagaaaa tgactaagaa gatgtactgt tcccacacag 120
tgttcaccct catgggatct gatctcaact tggcttccct gtgagaagaa aaacatggca 180
gcaaaaacaa gtatgatctt gtataccagc aatgggcaca tcaagaaatt atatagggcc 240
taagcaggta cctgaactga ggtttgctcc acaatgcagt cacagtaagg gagaatgagg 300
gctcatcagt ctatcacctt cccataagga aatgtaccct tctctaatcc taaagagcct 360
gtggcccata cagattcaa 379

<210> 43
<211> 1245
<212> DNA
<213> Homo sapiens

<400> 43
agacctgtga cacttctggg gagtcagtat gtaaaatagt tagagtgcag gctccagagt 60
cagtctccct ggttccactg ttcactagaa gtatatcctt ggatagatta cttagctcgt 120
ccattagatt tgctgtctgg acgtgtcagt agtaacaatc tcaatgtcac acggttattg 180
aagagatcca tcatagagct tacattctag taaggagaga caggcagtat acaaatacta 240
taaatacact ggaggaaaac atagtggggc tgattatagt cgatttttgc tcttttattt 300
ccggtaacag taaaacaact ctgataagaa aaggtttgaa catgaaggca gacctgaagc 360
gagtgaggga gtgagctatg cagctaacca aggaagagtt ccaggtagtg gggatagcaa 420
gtgcaaaggc actgaggaaa gggcatcatg ggaggggggca gcatgaaggc ctttgtttcc 480
agagctgaga gtgagggaaa tggagaactg gttttacaga taacgagtga tcggcatgta 540
tagggtatta cccataagaa ggactttgga ttttatcctg tgcgagatgg gaaaccattg 600
gagtgccaga gtaatatttt cttaaagcta gtctgcttgt catttctttc tttagaagag 660
cctcagatcc ctgttgcctt tggaataatg ttcacgtact ttactttggc acacaaggtc 720
cttcacagtc aggcctctgc ctgtttattt agcatcatct gcttctttcc cacatgtacc 780
cttcacttta gccaagtggg atcacatgca gctccctgga tgggtcatga tgctttgtgt 840
cttcgcgtct ttctgtaccg ttagccctgt gagaagccct ccccctcagc tcacatggtc 900
actggctctg ttttagaggg tcctctttgt gctcttgctc tctcatcatt tccacccgga 960
gccactttgc acttaagctg cttaggcttg aaacgtgctg tgttctttta ccagccttga 1020
ttttactctt gccacaccct tgccggttga aagtatgcct atccagctca gacactgcca 1080
tcttctgcct ccttgggacc ctcagaaaca ctttgtagct ctcttctacc taataatatt 1140
tatttggaat taaaatgttt ctttgccttc acaaaaaatg ccttgggcta acttatttca 1200
gttttctttg tagctcctag tacaggtctt acagacattt catat 1245

<210> 44
<211> 624
<212> DNA
<213> Homo sapiens

<400> 44
acaagtccaa caagacagaa agaaatgggt cagatgacgg caacactttt caaggccagc 60
caaatcttag gaggaacggt gaatctgcaa ggatgaagaa gaaacctcag ctacgaagag 120
tagttccaca tggtcacgtg gctgtgggtg aaagcactct catatacaga gttgaagaga 180
agagagcacc agggatccac caggcaactg cgttacagaa agaaagtcac gcacaggaaa 240
agcagatttc tgattctgcc accaggaagg gacatgaagc actgaagaaa agaccaatgc 300
ccgaagcttt caaattcaaa gtagactgtg ctaaattctc tatgatagtc accagcattc 360
agagaaagaa aaaaccactt ccaactgaac ttacctagat atttggtagg gaacggcagg 420
ggccgcatgt gagggggggt ttgcagaatg ggaactcaaa acagaaaggg cagcagaagt 480

cctcactaga aacaggcaac agccactagt ccactatgtc taaatccctg cccacatcgg 540
agaaaaagct taactcatca ataaaaacag aaagagaaac tgtcatgaag caacactcat 600
tttgggattt tataggacaa cgaa 624

<210> 45
<211> 801
<212> DNA
<213> Homo sapiens

<400> 45
gcccatcaga ggcattgttc ctttctgtta cattgttttt gattgctacc atttcctttt 60
tattcttagg atttccgtct ctctgctgac attgcccatc tcttcttaca tgctggctac 120
tttatccatt ggaactctta gcatgctaat catagttgtt tcggattcct ggccttacag 180
ttccaacagt cctgctatgt ttggttctga tgctggcttt atcccttcaa attgtatttt 240
tgccttttag tatggcttgt aatttttctt aataactgga catgaagtac ttattaaaag 300
taatggcata aataggtcct taggaattta ctggtaaagt gtaggaggag aggcagtgtt 360
ctacaggcct atgattaggt gtctgtcact taaagagcct atacgtctga actgtgggct 420
tcacaaggga ttatcagcac ctcctattcc caactgaggt ggcacaggct agtagattag 480
gctggagtta tgtatttcc ttccccaatg cggaaggctg gaattgggca tttctgtttg 540
tctaggtcag ttaggctcta ataaaacctc agtgggtagg ctgtggtgaa atagcttttc 600
ttaagggcag atcttgttaa gaaagttaga gtactctggg atacccaaaa tggtatcttt 660
ttccctccct ctgctggatg tacaaggtga ttttaatctg aaatttattg tgaagaccat 720
gcttgagctc tgaagttaaa atcaaatcac aaaaacaccc tctgcaccat acccccaccg 780
catgactgcg ttccctggag a 801

<210> 46
<211> 1027
<212> DNA
<213> Homo sapiens

<400> 46
ttcctgttcg acgatttcgc ggcttgatga attaccacac agtgaacaca cctgtgtacc 60
caccacctgg atctgacacc cagacgctgg caggcctccc ccaattacca tgcgcttcat 120
ctccccaaag gcgccctgtc tgcatttctc cacggatcat acgcctgtcg gcccttgtgc 180
ctggcttcac gctatatcac ttctgagagt cattcgtgtt gcctgtagca gtgggctgtt 240
ctttttatgc tgtgcaagtg tcaaccgtag gaacctgcca ccatttatgc ccatgctctg 300
gggctgctcc ccgcctggcg ctgtgataac acagccttga ccttcattac tggggaataa 360
aagtaacagg catatggccc aatctgtgca aaatattcta aatgttggca tttggtacgg 420
tctggtgccc caggccacac ctggatcacg accgaactac ctgcggaagc acctggtgct 480
gcctcctgag agacctctag cttaggagtc ttgtttttcc gaccagaccc tcagtcgcct 540
aaccacatag tgggtcccat gacgccagac caccagctgt actggcttga aggctagact 600
cgagctctca tgctgagagt gtgatgtgga cctctggcct atgaacccaa cctggccctg 660
gcttcaccct gcaacgtggc cattgcagac accacttgcc ttgcccctct tgggtacagg 720
ctcttccccc atgcccatct tcaggtggag acctcctgtg catcttctaa gcatggctca 780
gtgtcccctc catctttgtc tgtggtggca gagggagaca taaggccaga gggtgcaacc 840
gaggagggca tggagctcag ctctggccct cagcagcgct gtggcctcac acaagtcagt 900
ggcctctgag ttctgcatgg tgtggcaggt cttaaccatg gctgaccgct accaacaact 960
cctcacgtgt gtgccttgat agaagcgctc ttaataacca cagcctcact tgttacaccc 1020
atgtatt 1027

<210> 47
<211> 738
<212> DNA
<213> Homo sapiens

<400> 47
cttatttacc ctgtgcctat catgtagcaa gcacttcata aatgttagct attattactg 60
ccaaataatt ataactttac ttgggtcaac ttcttcacag aatgatataa aggagtatct 120
aacattctcc atccagttca aaccatgatt cttgttagcc ttttgatctt gatagtagag 180
cccctgtttg cctctctgac tcctctgagt ctctgttttg agtgtgttgt atttctaaat 240
gtaggccaac atcttactga ccaaactttc tctttaaatg ggctgttgtt tttaagcaat 300
agttgattgg gagggtgggg tgcgagttgt ggaatactta caagcccggg ctctggactt 360
gagaaaggcc tgagtttgaa aacttgactt gcttttcact atttgatcaa aggcaagtta 420
ttctctctcc agtgttttcc tctgaaaaat gggaatagta atatcttcat agggttgctg 480
tcagaattaa gtgaagtaat gcatgtgaaa tgtttgacat gatactgaca cttcacaagg 540
gttaaataga tattagctgc tgttggttta taacattaat atccatggtg gctgagactg 600
tggcaatgtt attcattgtg tatttctagt gtagagcaag tgatcagtat taaatgttga 660
atgaatgaat acgttacaga aaacaggcca tggtccattc ccatttactg tactgagtca 720
gaagatatac aagccagg 738

<210> 48
<211> 467
<212> DNA
<213> Homo sapiens

<400> 48
cgactgcact attaagtaga ccatgtctcc ccttctccca ctgtcctata aattggtcct 60
gtgtttccca accccaaatg gggtagtaac tcatggagaa caaaatgcca gtagtacaga 120
tatagaacat gggttaaaaa ctattctgat caaaccccca gcaagaatct tgaagaggaa 180
gacagagggg gaagaaagca ataggcttac tctccctaca acttaattcc atgctctaat 240
tcagagcagt ataattaacc ccgagtttca ctgctaacac tgaagttttt cagcctgatg 300
agaatatcat ttagtatcct gcaatcgcca ataatagcaa tatgacacaa tggtgcccat 360
tgaggctcta atcagagcac atagtcattt acatagaaca ttttctgtga ctctctgtaa 420
tacaacgcca caaagactca attgagacac aagtctcccc ttcccat 467

<210> 49
<211> 687
<212> DNA
<213> Homo sapiens

<400> 49
tcccggttcg acgatttcgt cctcagcctc cttgcagcag gagccacagc atcccactgg 60
gacccttaag cagcctgggc acaggatggt gatgggtctg tcctcacttg ggccctgttg 120
cccatgaccc acctgtccca gctgtgtgtg tcagtgcgtc tgcctctctg ccagatgttg 180
gagtctccca gtatgcccac tggactatct ttcagaattc cacaggggca gatgctctta 240
atgccccacc tgagtccccc tggcatttct tttctgtttc aggacacatg ggccctgtgt 300
ctggctgctg gcacatgagc ctctgcttga gggtttattt ggcactggac cctgctcacc 360
aggaattgat gcctcctggg agtaccctcc aaccaattac tctaagaatt ggcattgaga 420
tcctccagcc tcctaccctt gaagtgggca actctgaggc actctcggtg ccgtctcgca 480
caactcccag gaggactgag ctcccgtggc ccacggagct aactggcttt ctaataaaca 540
ctttatgatc cttcttccct tccctgactc acttccctct ccctcctacc catattcaat 600
gctttcctgg gggcacgatc tcaaataaga ttggatcctt ggctcaaggt ttgcttctta 660
gggaacccaa accgacgtgt attgcgg 687

<210> 50
<211> 457
<212> DNA
<213> Homo sapiens

<400> 50

ctgttgtggc tgaagtacat ttaaatattc agctcgccgc aactcataag ttgtgtatct 60
tccacagtcg acatactgat acctctcctg ggatatctgt tcagaccgcg tcgacttcta 120
ttataaaatg catacgtgtg ttccagagcg actattccag tgtcaccatt taatcagaat 180
cacatgcctg ttcatgattc tagaattcag gttgtttaaa tacggcagta acgtatgcag 240
tcatgtgatc attaatcacc cacaagttca gggacgccat agatgatcca gaggtaggca 300
tctgggtatt tcaacttctc gaggcccata tggaagttaa agagttaatg tcggctgaac 360
agaagattct ggcagagcac tgacactgaa tccaactttc acgcatcctc tctccacatc 420
attctttacg ccttaactct ttatgtcgac gcggccg 457

<210> 51
<211> 1422
<212> DNA
<213> Homo sapiens

<400> 51
atgggtcagg ttacttgtag ccaagtaaat gcttctcaca cctatgacct cacttcagcc 60
cctcatatga acaacagtgc acagccctat gcgcacagcc aaggcagcac ctgtgtagcg 120
tcattttcca ggctgatgtc tggaattccc cttagtaata tttctgccat tttttatgaa 180
accatcttcc ttgatggcca gttccttgcc tcccaggaaa ctgatgagga ggcaagacga 240
atgtattggg aaagcccagc ggaggaagag ctcctgagga gtgagtggcg tagctatctg 300
cttcatgtgg ggcgctcctg ctctgcagtg tatagtattt tttaggtgga ctaggagcaa 360
atgcctgcct gatacaggca acgtttgcac aaaaacccaa aggaagaagg ctgctggtag 420
gctaggggtg gcaggtggga ttgcccttgg gctttaaaga aggctgagaa aacttacttc 480
cagagtggga ggcaggagga gtgagccttg gaacaatagt ggagtgtttg tggtggacgg 540
acgggagaga ggaacacttg agggcatgaa cttacttctg tctctttact tgaaaaagca 600
atccactgag ctaagaccct aaaatacatc aacccagagg gcttcctttc gagtttttag 660
tttttagatc attcggtatc ctcaatgcaa agatgtgata aagtcaagct caggactact 720
tggggaccaa ttctcatcct cctcatgtgt tacaattcct tagcagttgc tttaaaacta 780
acatttttct aatcttgaaa tatttaataa agcagctttc tgctaaagag ctgaaagcat 840
tttacgaacc ttatcttgag gatatccgtg cgtattttat tattttaagt catttttttt 900
agcctacaga atttcatcaa cttctttcct tcagttacac gtaaagtagc ctcccttccc 960
ccattttata gaagaaaaaa caaatcttca taatgtttag ataattcacc caaggtcact 1020
tggcagaatc tggaatcaaa cgcaggtctg ctggttccaa aaccttctcc cctgtggcat 1080
taggtatatc acatggactc ttaagtaagc ctcaaaatct acaaaagtgc agtgatctgc 1140
ccagactggc cagccacaat gtcaggggggg aagggtggga taacccacat gcccggagac 1200
aatggcggag acaaacaggg atcttctctc ttctgagtcc aaaacaaaac ttggggtctc 1260
cctggggggc acccacagcg gcccaccccc ttcgtgtttc accagcatgg gcggaagtgg 1320
cagcttcttg ttctcgtggc gggagcggac acagagcgta ccgacactat ctttttctaa 1380
caacggtcct tttgtggcgc gagcggcggc caacgaagcc ct 1422

<210> 52
<211> 447
<212> DNA
<213> Homo sapiens

<400> 52
aattctcggg tcgacgattt cgtattaaga taaaattctt tcatcatgtc tgtgccctct 60
aaaataatac tcctattgct ttgtcttccc cttctccccc cctagtcctt ctttagtcat 120
tctatttaga atcaagtcgc tcatgagttt aagaattaga gcagcaagaa attgggctag 180
agatgtacaa aagctttgga caatagtagt tttgcttgtc ctcattctta ttagaagtgc 240
tgttaattta ctgataaatt ctaggacgga agacaaatct ttgcaactgg tactatatca 300
gtcggtaatt atttgttttc cttagaccat attgagctcc ccaagtacaa gaacaatgag 360
tagatgtgaa ccaaaatgcc gggatttttt aagtttccaa tatgtgtgtg tatatataag 420
agacagaatc ttgctctgtt gtccagg 447

<210> 53

<211> 367
<212> DNA
<213> Homo sapiens

<400> 53
ttcgtgtcga cgatttcgtc gacgctccct atgccccagg attaagtcca aacaccatgg 60
tgtggcatgt gagaaagtct tcctttgtct ggcttctgca gctcttcagc ttcatctctt 120
gccactctgt catctctgtg tccccagtgc atgtcccatg gacacagtgt gcagtcatac 180
ccccatacac cagctgtcct aaactatttg caatacaggg aggaaggttc tagatagcca 240
gatggcttca ttccgtagct ttgttccttc cctctgtttg ggatcctctt gacaagcaca 300
gaaccatctt tctaggctcc attaaagcat acctttttctg aaagaccctt cctggtcgac 360
gcggccg 367

<210> 54
<211> 917
<212> DNA
<213> Homo sapiens

<400> 54
tcccgtgtcg acgatttcgt aagatcttaa ttagctttta tttgagattc taaagttggg 60
caacacctca ttcttagaga atgagggttt caatgagctg agcagaggaa gttggtttta 120
tagacagggg ctgaggaaag cagaaacaga aaacaaaaaa tgggtggttt caaagttaat 180
ttatttgtaa aggttaaagc agagggaagt cccttatgct ggctaaaact ggcatgtttg 240
ggggcttggc tattatctct tctaatttct cagaagtcag atgaacaact ttgctttggc 300
ttgctgacat ggaactttag cacaagtgac tccattttgg tttggtttgt tgggcctagg 360
gcagggactc agtccaaacc agtggtctcc tataaatttt aacagatcta ttagtacttg 420
aaagaaatgt caaatttcat ttagaagtta acaaaaatgg aagatgtagc tgttcaggcc 480
cctttaaatt catcccatgt tcccagaccc ctttaaagtc tatccatgga ccgcaggcta 540
agactgcttg gcttaggtgg agcagccagg agaactgcat ccccccggat aggggatcgt 600
aaaggcatta gaggagggag gcacttgcct ggcagagggg caaactccaa aaaaaggcaa 660
agatatctcc cctaagtgtg aaattttaaa aaaatcaata gctaacaaat gaagctggaa 720
aaaccacttg agacaattgg gaaaatatta atgctgaaac gtaatagatg ttaaggaatt 780
attgagttct tttcccttag gagtggtggt taaaataaaa cgtcctagaa atgtatgaag 840
atgtatactt gagatgcctt ggattagctt taaaatggtt ggaggacgag gaaagtagga 900
gagtagagat gaaataa 917

<210> 55
<211> 688
<212> DNA
<213> Homo sapiens

<400> 55
cgaggatcca aatacgcggc ccgcgtcgac agtttccatg acattttcac tgcttttaca 60
aacccactct tagtatgttt tcttgcatta ctacattcag taatgaatat aacatactaa 120
gagtgggttt gaaaatgtca tggaaacttc ctctgctttt acaaacccac tcttagtatg 180
ttttcttgca ttactacatt cagtactgaa cataacatat actcctccaa aaaagaaaaa 240
cgaaaattgc tccaaacccc tgatcctgac cagttcccta gggactgtac aataacgtta 300
caagcagcat taacactacc accggcaatt aaaaaacaaa ttaacccaat ataacgacct 360
cagaccatgc tgcagtaatg acagtggcct agcagagcga gcgagcaggc tgagggatgc 420
atacctcctg gagtttgggg caaggaaaag gagagccaaa atcatggaaa catgcaaatg 480
agggagaaaa gaaaattaaa acaaagaaat taaacccgat tccttaatct taaagtgcaa 540
tgtatcagtc acattcatgc ctgcctgggg agagagctga actaaaaaca ctcaacgggc 600
aaatgatccc actactgtct ggcacgtgaa cgacgggctc ttacatgcca cgttccacct 660
catcaaggca agattgtcga cgcggccg 688

<210> 56
<211> 1003
<212> DNA
<213> Homo sapiens

<400> 56
cccgttctac gattatcggg catcatttca tttctcattt cttcttaggt tatgcgatct 60
acttaaaatt atatagggta gtgtcctcaa ggaagtctcc tcttccagga aaggaccagt 120
gtccactgag ggaccttctt caggacatgc cccaggtcag ttgttggatt aggcaggcct 180
tgggcgagtg aggcttagta tggtttggga cttgagtttt tcttgctttc ccagaattca 240
tgagcttttg tcatttagaa tgtttcttac tcttctgtag ttgtcctcac tggcaaatcc 300
cgccagatat tcttcagtga gtatgcttcc tgagtgaaat ataaaggtga aggaacattg 360
ggcagagcag tattaactac ccatggccac atttttgtta attattttgg tatttattcc 420
ttgtgtgggt gtgtaatttt tggtaaagtt gcattaaaat agtatagcat gtgcgcattt 480
ttacttcctt ttaaacttct tttcttcttg gaaatctctt tggctatgaa atctcatttt 540
cctttcaccc ttctgatgct tagccgagtt ttacttaaga aaacggtgga tggcttgaaa 600
ttaggttggg taattacaat tccttcgaat gaattgactt ctgtattcac cctgatgata 660
aatagaccaa atcaaaagca tttttgacaa agtaatgctt tctcacatat ttagcttctt 720
tatagacact tacccatctt gttaaaaacc caaactagtg attttacgtt gtaaaaagcg 780
tttttcaaag catgtgggcc atgtaggaat gagacgacag gggtaaagaa ctatgccttt 840
tttgaaggga gaaccaaggg gttgggtttg aactttggaa tttctatttg gatttccaag 900
ggaacctttt atggaaccag ctttcattaa attctggtta aaaatatttc acctttccg 960
gttaaaaaca aaaaaaaccc caaaactcat tgtattgcgc ccg 1003

<210> 57
<211> 401
<212> DNA
<213> Homo sapiens

<400> 57
tttgagtcta ctaacagtag actattagtt agggagaact tactgaatgc tcactctgtg 60
cgagactctg tcctgaatac tttttcctgt gagatagatg tgcttagtat tttcacttgg 120
cagatgagga ccctggagaa caaagggtat ggaactttcc aggcccccac aatgtgtaaa 180
gcacaagtgc aggattcaag ctgcagagtc tgagtcccaa aggaaagcat ccaggacatt 240
acctggagca gagtgggcat ggagtaagtg acagtggtga gttttggtta ttctgtctcc 300
acatttaacc aggcttcgct gtagcctaaa catcccacct tcttggtgat ttcatacaca 360
cattcacctc ttgttgaaaa gaaagcgtga gaacctaaga a 401

<210> 58
<211> 206
<212> DNA
<213> Homo sapiens

<400> 58
cggccgcgtc gacagcatgt gtcactaaat gtcactgaac atttgtgagc ttgactggaa 60
ccatgttgta ggacctatat ctgatggtgg tatgaggatg gactggaaga ggcagccact 120
ggaaagagct gttaggggag cacaggagaa gtaataaagc cctgcacatg gtggcctcaa 180
tgaatataga gaggagggaa cagatg 206

<210> 59
<211> 649
<212> DNA
<213> Homo sapiens

<400> 59
aaacattctt tccctgggaa gtttttatt gatccttgtt tgaagttgat acataatatt 60
atatcataac tgtgtgatgc tgtttctgca tatagctgaa tgcagctttt tgagactaaa 120
ggttgcattt ccttcatctt tgaatttcca gcccctagca cagttcctgg cacatatttt 180
ggaggtattt tataaatgcc tttggaaaaa aggagttcaa gtgtttaatt tcctggcaaa 240
ctaaatagct ataaagagag agcttagaaa gtctagagta agagatgatt agcatgcttc 300
ctagaccatc tgtcactaat accaattgtt tctacttggc actgcagact tgtttttga 360
gggggagggc acgcttgtgt aaagttaaaa aaaacagaac aaaaccaaat accagaactc 420
ttaggttatt ctatacccct ctgtttagtg gctaaaaagt cattgctgat ttctaaaaaa 480
catggacatt cttattcagg agagatagat gccattattg ctttttccaa gagcaataca 540
gttgctataa ctgagagctg gttctgcctt gccccccttac aatctattct tgatgcggca 600
accccgagtg atcccattca aaggtgcaag ggcaagaaat atgtttctc 649

<210> 60
<211> 408
<212> DNA
<213> Homo sapiens

<400> 60
aactttccag aagtccccca acccatgcct ttagatactg aggccatctt acataaaaca 60
gctgaatggt atgtcctctg tctcaaaacc tgtttcttca tgtatgtcct ctatttccca 120
tatttgagga gtatgagtct gcttgagtaa cagcacctcc atccttttga gagaaataga 180
gaagtggacg ctgggtaagt aatgagtcta gttttggatt tactgaattt aagtcctata 240
tgtgactaac tggatactca tatccataga cgcatttact agacatgtgc acgtagaacc 300
tatgagactg ttttgcgact gatgtctgat atttaggagg cctgtgctta cgttgtggta 360
catgagtatg ccaaggataa tcttgtagac ggataaatag tcgaggat 408

<210> 61
<211> 756
<212> DNA
<213> Homo sapiens

<400> 61
cggccgcgtc gacccttggc tgacctataa cactgaatca catgcagtag tatttagttg 60
attctaccgc taaaatgtca aacaaatttg cccagtttta tccaaccctc ctattattgg 120
ccttcattat ctttcaccaa gaccaggtaa attatctaat ttctcttggc cagcttctct 180
tgctccaacc aaaggacaac tttggttgga gcaatttaag caaaactcaa attgtttaac 240
tgccttgatt ctataccttg catgtgcaaa tcaaagtatc tttgtattag acatcagttt 300
ccttattagt aaggtgggac aataatattc catgagattg taaaaagttc tgatcataag 360
aaacgtcagc agtaatataa tatagcaatt aataaatgcc cccttacttc ttgtcgccat 420
tatctcactg gggaatatta taaaacctat catatatggt gataagggtt atatgttgat 480
taaataattt agtcttctca acagctttat gaggtaagta tcactgctac aattccaact 540
ttagagatga ggcaaagaga ggctattttg tctaagatta cacatcaaga caaaggcaac 600
actgaatctc agatttaggt atgcctgagt ccagactgtg ataaccaaca gaatatacca 660
actctcaaac atattaaatt ctcattacaa tggaattaac ttcatcacct tcaaactgtt 720
gaaacaagat tttaaataaa agaaaatatt aataag 756

<210> 62
<211> 675
<212> DNA
<213> Homo sapiens

<400> 62
tccggtcgac gacttcgtcg aaactgaaaa ggtagtgaca tatggtggca ggtgtgggag 60
gtaaggaggg aaatgaaaaa actggagttc tccatgaggt tgggggggctg ggtggtgggg 120

atggagctga gccaaagcca caaagtgaat ggtccaagct cacagtttct aggcagcgtg 180
gtgcagctga tgacaagatc ctgggcatgg ctctttgaaa gggttatctc atgtggagca 240
caggtgaaaa ccactggaac tggggaattc aaggttaaca tgtggctagg acatctgagt 300
gttccccaga gctctcatga tgctgtctct gtggctggga agagctacag aggtttatta 360
gagaatgtgg aataaagagt taacgcctat agaatgatgt ggagagagga tccctcttcg 420
cggagctcat gctcgctttt atcacatcaa aaccaacccg ctcatcccct ctccggcccc 480
gtccgttgtg cgcacgccgc tgtctttgct cactctacac atcctcgccc cattcccact 540
ccccccttgc cagcgacgat ctattcaaac tttcgccctа tctcctttct gtcgtcttcc 600
tccggcccta cctcgtcacc accgctgtct ctcatccccc ctcctcttgt ctacactcca 660
tactggtcac tctcc 675

<210> 63
<211> 921
<212> DNA
<213> Homo sapiens

<400> 63
gcaattctcg ggtcgacgat ttcgtaaatg tgtcactgtt tggcaccaag ggcccagtga 60
aagggcagtc agttttgacg tttttaactt taagtaatgg gatgaaaggt attagagatt 120
taaatttttа atttaataaa taacatgaaa atcaaccacc tttcttagcc acatttctgg 180
gaatgaattc tattaaatac gtacgtaact actggccatt aactgtactt ttttcttcag 240
gcaaaatgtc atttgcatag accatattcc tgcagcagag aataaatgct attgagaaag 300
ggctttaaga ctgagtaatc aggtttctgt taatattaag aggttaaaaa tgattctagt 360
attcaatatg agcttccagc ctgaagttct atttaatttc aggttataac cttaaataaa 420
caattttcct agcctcatta atgcttaatc taatgttgca taacaaagta tggacttcct 480
tagtcggttg atgctactga gaatgtgcaa atgtgtcacg gccacatacc agtacattag 540
aagatctctg tttttaaacc ttgtccctct tctacaaacg ttgtctatcc tcactgctca 600
cagtgtgctc ctcaggccag cactctcatc actggttaaa atggaagatt ctcaggccct 660
gtccttgtct cttgagccag tatctgcaat tttaacggtc cccccagagg attccaaagt 720
actgtaaagt ttagaaagca ccgctattta cagtaatgca gacttgcaga attctgagcc 780
tgcatagtat atcattctta aaatgatact gtctttattc tgttaaaatt gtaaagtttc 840
actattaggg cttgattttg tttttctaat tattgcaaaa tatttcaaac agaaaagtat 900
cagggcttat ttccattgtc a 921

<210> 64
<211> 507
<212> DNA
<213> Homo sapiens

<400> 64
aattcctggg tcgacgattt cgcccgactt ggcctgaagg aagtgacatg agcataagga 60
agtggggtgc ggtgggagcg ggtcaagtga cagaggaggg ctgggaaata ttaatgaata 120
gagtctgttg agatccaagc caatgtgact taaagcacca ggcaaccaga atcgtgatgg 180
gtcagtattt cccccaggcc tatccagacc aggtcttatc ctcctagcaa cgcggccctc 240
tcggctctta aggcgtgact tgccatggat gttctgcctt tggaaccagt gggtggtcac 300
agcacagagg ctgttggttt catggctgtc acatgcacag agacagccct gtccactgag 360
cctgttctgt ggcagaagga acccccctagc ctggaccatc tttgggtgga aacaccagcc 420
cctcacgtct gactgtcatt ttcagatgtg aagagaacac acagtgttgc cgatgtgttt 480
gtattgcctg taaacaacga ttgtatt 507

<210> 65
<211> 391
<212> DNA
<213> Homo sapiens

<400> 65
ggcacgagga tcttttgttg cccttccttt ttaagccttt cctgtgcata gtgaagaatg 60
cgcatgacta ctagtagttt ggtgttaccg cctttatttg tgctaaagtg ccagcgtttt 120
tacccaccat tgtatctgca cccttacagc atatgtcaac atgttagtat tcttgtaaaa 180
aatagtttgg gacctggggg ttctaaggtc cccactttgg ggaaccattg aaataggtac 240
ttagatctac agtgtatcat atcttttcat tctacaagat taaaaacatg atttcagtta 300
attgtttttg taaatttctt aatatggtgt tgaggggttt cagtccagag tgaagacatg 360
tatgtttctt ggcttatgct gaagtttact t 391

<210> 66
<211> 390
<212> DNA
<213> Homo sapiens

<400> 66
tttcgtccct ctgtaaggac gttttgtggt ggagccaggc aggagatgca aaggttggtc 60
cccaggatgt gcaggaagga ggctgagtgg ggagggccag gaagccgaag gcctgagcag 120
gaatccggga catcaactct tatgctcgca aggaggcatt tgtgcccatg gggtcttggg 180
ctccctacaa ccgcagcacg atctggggag agaacggaga gaagggagag ggtgcgcaca 240
gccagccccc gtaaaatcct cttcaaaacg cagccaccca ggggttctag cacagacaga 300
tgcccatggg gcaggcagtg caagcatggg acaggcacct gccatatgcc caacagataa 360
caggacaact gtttacccat ccctggtatt 390

<210> 67
<211> 804
<212> DNA
<213> Homo sapiens

<400> 67
gaaaaataag atgtctacat atgtgctttc cttctactac tgaagactat ttaaaaacta 60
accttcactt aaaaaatact gaacatacaa ctaagtaaca ttaatgtcta ggtctaggtt 120
actaagttct cagtaaaaat gatgtttagt aaggaaatga tgtcagtcag cctgcagctc 180
atattgataa gccatccaaa attggtccca taggtgatga catctagacc tcacttcttc 240
cgctacctgt gctctcttcc ccctcttctc tttcctctgt tgtcacaaag ccagctactg 300
ccaggatccc ccctccccat cgctctgcag tccagggtgg ggtctctgtt agcatgagga 360
aacgtgaaag gaaatgtatc tctgtgtacc tttatttttcc ctccagggct taaatggaaa 420
aagagagaag gtggagtccc ttttgccatt gtaactggaa ttaaaaacct taaaaaactc 480
aagttggagc tctgagtgca ttttcccaca aaaaccatgt tatatatagt ggtttggttc 540
tatggtacta ttaataagac agaatacttc agctatgatt tctgggctgg cctgggaaac 600
cactaaatat gacatagttt cagtgaaaaa gtgcgttcag ggttcccagc aatgagtttt 660
acaaacccac aatcagaaca caacccatct gaaagctgag ggccgcctac ttcgcagtag 720
gtttgacagc ctcacaaagg catccgaatg tcatagacgt caacgccaag cttaagagtc 780
tcctgcgtgg cgctccagca gtac 804

<210> 68
<211> 386
<212> DNA
<213> Homo sapiens

<400> 68
tttccctgtt ctgatgcccc agccccgcgt cacagcccgt tctcagtctc ctctgccttc 60
tctagtagat attgttagcc aggattccat tatcctcaat gctcccttcc ttcctgccac 120
agtccttggg aaatctcatc cacacacttg gatttttgct gataattcac aaatatatgt 180
cagctttcaa aaaccgcaca gatgagttta tgaatatggg tatgcagccg tatataaaat 240
ctccttaccg actatcaatg tctcagatca gcctgaaatt tgacttgtct caaactgatc 300

ttatttttacc ccacaaattc tactctcctt cctccttccc tactgtcatg ttattctatt 360
cattcggtag gttaagccat aagcct 386

<210> 69
<211> 282
<212> DNA
<213> Homo sapiens

<400> 69
cgacgataaa acaggggaag gagggaacag gtagggctgg ctgcaggaaa gcttgcaggt 60
atggggtaca cagttaaaag ggagtgtctt ctgatagtct tgatgcaggc ctggaggtcg 120
tttgttatgg gggtggaggt cctgatgtac attgtggcgg taagatgccg tgctgatttt 180
gccaccagcc tgtggcagcc ctggtgttac accagggcag gtgggcagtt taatgtatct 240
caggccagat gacctgccct gcccacctgg tcgacgcggc cg 282

<210> 70
<211> 338
<212> DNA
<213> Homo sapiens

<400> 70
attttttttc atggttgtcc tatatattgt cagagcatat aatcattaca tactatgctg 60
tctctcctcc agtctttact tagtcttaat tctgttagta accgtatatt taatgctcac 120
cactagttct tataatgatg tttctttagt catttggatt gcctcaagtt ttgcctctag 180
taaattcttt aggaaaggct taagggaata ttcatatttc atgaattttc ttgctcgttc 240
ataacatttt gtctgtgttt tttatatcta aagtttgtgt ttgccagaaa tagaatcctt 300
ggctcacttt ttctttttctt gaatatctta aatttgaa 338

<210> 71
<211> 380
<212> DNA
<213> Homo sapiens

<400> 71
tccaccttca tgttaccctt ggtctcccca gtctctggca gatccttctg acctggcatc 60
accaagacaa atgctggagt ggcctctcct tggacaaatc cttcccatga tcatcccact 120
acctcctctc cctgccttag tggtttggcc cattgggctc actcactgcc cttggccctc 180
ccctttttgtt ccagcctctc tggatgggtt ctacaacagt cgcagcctgg atggtaagtc 240
tccacccttg aggccagaga agtggtctcc ctggtcctgg ttctcacggc tcccaggtca 300
tgggcttccc aaggagggag ggcagagaaa gagagaggtt tgggcctagc aggttaaaaa 360
gctccttcgt cgacgcggcg 380

<210> 72
<211> 391
<212> DNA
<213> Homo sapiens

<400> 72
ctttgaaagg gccacatgct ctgccatgct agctgaactg cataagaatg catcagttat 60
gcaaagcatc ttttaccaca attgcacaaa tactactttt tatttaaacc cttggctcat 120
gattaatgta ttcatctaac aaagtcccaa aataatggaa ggatgatgct gtggcaggtt 180
taccctggac catctgcgca ggagctctgc ctctttctcc atgcaccctg gtccatgtcc 240

acagcagtgg agagagataa gagacagaag gatggacgta gtcaccgcat gttgctgccg 300
cagccccaga gcctgatgtc tagatgttgc taggtagacg tccagagttt agcaggatag 360
attcttgata ttctgctaga atcaacaatt t 391

<210> 73
<211> 363
<212> DNA
<213> Homo sapiens

<400> 73
gttaatttct gattctgtga aggccagatc atagagctgg aaatgattga aaagaaaagc 60
atgttctttt gtttaacaag acaaagccta ttgtgtactc tgcttctgat gcttaaaaga 120
tgtatttttt tttcttactg cgtaatatgt agggctaaaa gttttgaact cttcacttca 180
gaaattactt ttccagacaa aagagcaaaa caatgtttgt ttaaattgtt ctctggaaca 240
tgactgatag tgtttacaac atttctatga cattttgatc ctgtggggag tacttggaga 300
ggtatttaaa gtgcttttca ctaaatcaga tgctttgtag tgtttgaacc ttttttaaaa 360
aaa 363

<210> 74
<211> 1120
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(1120)
<223> n = a,t,c or g

<400> 74
tttcgttgaa ttattggcca cttggctggt aattacaaat ggtgtgggtg gtaggcagtg 60
gctgtcagtg ttcagtgctt taatgccatt gtgtaatgga tcagtaatga accttaaggc 120
tctttctggg tttgtgtgat aaaccttttt atggctacac agattttctt actgatgtag 180
caaagattgt aactgttctc attttgaaag caaagatgcc catacaacat gggaaatact 240
gatnnnnnnn nnnnnnnnnn nnnnnnnnnc ttcagttatg agttggttgc tggtaaaact 300
aaagcccagt ttggggttcc ctttgcagaa ttttctgnnn nnnnnnnnnn nnnnnnnntt 360
actgcttgtc gttttctgta tatttaactg gagctgggga ataattggtg tttggcaatc 420
tctgtatcat tcattgtagg caaagctgct actgtcaccc acagatacat ttgtttgcat 480
tccattgcaa ggattgaagt cagcacattc ttagtcccaa attgcttttg gatattaatg 540
tttatcactt ctataaccat aaggagaatc aaaaggctat tggaaaaagt caaagcttgt 600
tatgcgcctt acatattggg tccgagaacc cattaaaatg cccgaggtct aaaatacgct 660
ccagaagttg cactatcatt tcttccttct ttcttccccc ggcaacgcta tttcatttgt 720
cccctcccgc gtcctgccgc agttcacaaa cactccttcc ctgcgcctcc agcgccgcca 780
tgcacccata ctctccatcc cttctcgtgc cgggtcgtgg caacacacat ctcgcgtnnn 840
nnnnnnngcg ttccgtcccc ccccacacac tcccctctcc ccctcactcc cggtatccag 900
ccctccccac actacaccgc tggggcccgg cgccccgccc ccccccctct cccgtctcgg 960
ccgcggcccc acctggacac cacacaccct ctcgcccgcc ccccccctcc tcactttgcg 1020
caaaacccca catacccccg ctccctcgcc cccggatcac agcccgatca ccaactcacc 1080
ccgcgcgccc ctccgcgtcg ccccgcgcgc cacaccgccc 1120

<210> 75
<211> 353
<212> DNA
<213> Homo sapiens

<400> 75

cgcggccgcg tcgacctagg cagagaaacc atgagctgga taatctttcc ttgtctttga 60
agtctctgtt actgagaact ggtgagaaca ggtgagcaaa gacaaatgaa tctctgccta 120
agagcctatt gctgttgggg actacatgtc catattggtt gtttctgcct ttctggctaa 180
tttgaggctg ttgatgacca tctcgacttc ccagatgtta aatatgacaa aaataaccta 240
tcttgtcctt tttcttcacc tttcagctct cagaattggc agtacacccc acagcttctt 300
gctaaagtcc tatcacctgg ggacccactt ttcattattt cacatgaact cac 353

<210> 76
<211> 341
<212> DNA
<213> Homo sapiens

<400> 76
tttctctcag ttgggctttc tcatcacctg tgggaaagct gtgttggtaa ctttattcct 60
tccaatgttt gtatttgtct ccatccatac cgagcttgtt cccattttaa gacctttgtg 120
cttgctgtac tgctgtccag attgctctgt acccagacct ttgtatagcc tcaaatatct 180
cctgcttgag aaatgacttt cctgaactcc aaattcacat gtcctccttt acccagtcac 240
tccattatat cattctgtct tatttctttc atagcatttg tcacatcttg cttgtattgt 300
gaaattaatt tacttgctta atgtctataa ccatctctaa g 341

<210> 77
<211> 441
<212> DNA
<213> Homo sapiens

<400> 77
ggaattcccg ggtcgacgat ttcgtctttt tcttatcata ccacagcatc ttggatttac 60
agcaaaatgt taaatagaca tgatgagagt gatcatccta atctggtttc gaatctcaaa 120
gggaacattt cagcattcca ccactaagtg tgatgtttgc tttagagttt ttcttttatc 180
aaactgttca tttctaagtc ttaattataa attgacttca gatttcatca tctattgagg 240
cactatgtga tatttttcct ttacttttaa aatcctttaa gacattatta ctttacttt 300
taatgtggca attatagtac attagttttc tgaatggtca gccaaccttg caattactgg 360
aataaaccca attttatcat gttggattat ttttatatat cgatggaggg ggtttgctaa 420
tattttctta ggaattttgg c 441

<210> 78
<211> 342
<212> DNA
<213> Homo sapiens

<400> 78
tttccctgta tatagaaaac ggatttaaag acaactgttt aagtacctct ttaatatcaa 60
aacactaaat ttacaaattg cacatttatg caaatggttg taccaaggct tctctctgta 120
ccccaattgc tcaacacagc tccactgttc ttaccatggg agaaaaccgt aaaaactcag 180
tattctggaa taatctttaa attcaaaagt agaatagaaa cagcagagaa aagtataggg 240
gataccaaag agagaattca gccaagtcaa atatagaata cactgtcttc ctgcttttcc 300
tacctaaacc acattattgc cctattctta gtcccccaa aa 342

<210> 79
<211> 390
<212> DNA
<213> Homo sapiens

<400> 79
atggggaatg tgatgttctc ttgggtatta tgctaatcat attttggcag gtttccctgg 60
agcagatgca gaaatgatca taccactttc cagggtgtat tattttagct ccttagactt 120
tgtccatgct ctgtcctccc ctccccactg taaatactct tcttttgggc taggggctg 180
tgccctccag ggtattgttt tccttcctgt tcagtgccca tgctcttgcc tgtgtttctc 240
ctgtaccttt cccaggacct agcagacagc cgagcacctg cccactgctc tgtaaatact 300
gacctgcatc tcaagtgggg tagcctttgt gtcctttccc atttccaagt tgatttgcct 360
gtgaatccca tatgtgtaca catatgcaga 390

<210> 80
<211> 323
<212> DNA
<213> Homo sapiens

<400> 80
atttgagtca gttcagccct ttgagaaacc taggatcaag agttaataga ccatgaatct 60
agaacatgtc attgtatctc tcttaacatt ttatagagtc ctcctctata aagagataat 120
tggactgcat cattgttttc aacattttca tgttaatgca tttttactat caccctccc 180
cccatcatga ctattgccct actttgagta gacatagatt gagaatatgt tttggctaaa 240
accatcacat tcatcagtta cgagccactg gttactattt cttaaaggaa attttgagaa 300
aattgtgaat gacagcactt ctc 323

<210> 81
<211> 383
<212> DNA
<213> Homo sapiens

<400> 81
atacagggtg gggggagcac actagccctc aggggtacac acgtctgcac aggagtggat 60
gcacacctct gggtgcacat acctggtata cacatgtgca agggcatggg tacacttgct 120
cgtgtggata tgatcacagc tgagtacatg ggcttgtgcg tgcacaagaa tagaaaagga 180
tgtgcatgtt tgtgcatgca cgaggacaca cagagcacac gagcacccac aggacacata 240
gacaagccaa agggcactca tgtgtggggg tcaaacatct gagagcgcgt gaccagtgtg 300
agtctgaaca ccatacactc tattgacagg tgtctgatat ggccatggtc tgcactataa 360
ccactaagag gatcactggc tcg 383

<210> 82
<211> 214
<212> DNA
<213> Homo sapiens

<400> 82
ttaaaagtga tgaagaagaa agaaagaggt tatgtcctag aaacacagat acttcttggg 60
tttcaaatac ttgctgtgaa caaattgact tctctggtga atgcatacgc cttccctgac 120
ttcaatcaaa gaatagaatt tcatacactt aaaggtttgg gggaaaaaaa acaacaaaca 180
aaaccaaaat ctctgggaaa tgggaaaaaa gaag 214

<210> 83
<211> 94
<212> DNA
<213> Homo sapiens

<400> 83
gaaaattcac agaagcctcc ctggtaacat tttgctgcac tatgatagaa cttgcatgga 60
aattctaata tgcatataaa tgcagttctc tcct 94

<210> 84
<211> 358
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(358)
<223> n = a,t,c or g

<400> 84
atcactaaga agatgaatta acttagtctc tgatatttat ctaagatgga ctcctgtcaa 60
aaatctctaa ccagtcctcc atgcccaacc gcatagaact ctctgctctg ttcagggtac 120
catttcatgt tgattggctt nnnnnnnnnn nnnnnnnnnn nnnnnngtgt tatgtccttg 180
ggttctgcct accttccaca gtcaaaacac attcgttgat tgtgggtgct gtttttcaaa 240
gtgagattgt cactaatgtt tgccagattt ataaatgtaa aacaacttga cagattttgt 300
atgagttttt tgggtgcaga tgacttaagt tctggaatta ttaacatgtg actgagag 358

<210> 85
<211> 1092
<212> DNA
<213> Homo sapiens

<400> 85
tattgcttgt tcgacgattt cgctgcacag actctttgct gcatggctca ctacctgtgg 60
tgttggaggg gtggttttga tgcttcagga ctgtctttcc tgttctcctc aatgtctctt 120
tcaagatatg aagttaaaat taggtactgg gattgctcac ctgagttttg gtccttgaga 180
tgatgctttt ctgtgtgcag atagttgcta aaatttggtg ttcctgctgg gggtacaaac 240
tatgtaggtt tctattctgc cattttgttc tatcctttat caactattta ttttaggaaa 300
aggaatgtgt gtacaattgg agagccattt acaaaatctt tactacaaat gtttttgatt 360
ttataatttg tgataacagg aactatactt ttataactct ataaacctct gggtgtccct 420
aatacataga atagtgagag ctcgagagtt agttttagtt gcctttccag gtaggaccat 480
cagaaggaat gacagagcag tttgttttga aggtagacta aagcaccaag gtctaggtgt 540
cacggaaatt acgttttact ggatcgagtc ttgctctgct cggtctttcc cttttaccct 600
acatttcctc ctctcgatct tccgtgcagc gtaccacgcc catctgactc gagagctccc 660
cacctgcaaa cctttatctt ccttgttgcg gctgcctagc ggatcccctc gccaccgagg 720
ccagattctg cttgttcaga acttctgccc gtcttccgtc ttcttttaac gacgcacacc 780
cggccacgcg cggacggacc tatcccctcg ccccggtcca ccggggcggc gctggactta 840
caccttctc ccgcgctcct ggggccatgc gcgcgcatcc ccataccca cgctccgccg 900
cgcacattat actctctact tggcttgcgc tcgggcccta tctctatcgc actcgtaact 960
gcggtccctt gactgaagaa ccatccctcg caccttctac tcagggcgcg cggcgtacta 1020
acactactcc cgagctcgcg gcccgtccgg tcgctccgtg ccttcatata tcttacacgt 1080
tcacgcgcgg cc 1092

<210> 86
<211> 415
<212> DNA
<213> Homo sapiens

<400> 86

agtacattgc ccttgtctgt atacagtatt atactttatt ttaaaagttt acaccaaata 60
atcgattcat tctctcagct tactataaca ctcaaccatg gaaattctta tctccactgg 120
ggcctttcag acactctcct gtttccttcc tctgactctt gctggtttac ttcctgctat 180
aatcctctaa atatggactt gacccagagc tcagcacaca attaactgct cttgctttac 240
attatatgcc tttgctttac attatatgcc ttagacagct tgttctgttc cattccaaat 300
gccactctca acactcctgc agagcgggag gcatacagta ctcaatgcac gtgagtctct 360
ttctttcttc acccataaac tatgacaatg gttttttagt gtctcccacc ttccg 415

<210> 87
<211> 451
<212> DNA
<213> Homo sapiens

<400> 87
aattcctggg tcgacgattt cgtaagatcg atagtatctc aggcaagtaa agtggcaatg 60
agatacagca gagtatacag caacttttac atatgtattt tggaggtgga atcagaagga 120
tttgcttaat gggttggctg tggagggaga gggaggaatc aagggtgcct tcaggttgca 180
ggctccagcc ttgggctgct gatggtacca tttgctggtt ggagaagact ggaggaggaa 240
caggtggaga ggagtgggta gcaagagttc tgtgtaaata cttgggaggc atccaagcgg 300
agagttaagt aggcactgaa tatttaagtt gagctgaggg gagtgatcta cactggacat 360
aaattttggg agtcactagt atacagatgg catgtcatgg aactgattga gattgtttgt 420
ggccttaaga tcaagccctg cgagactgga g 451

<210> 88
<211> 74
<212> PRT
<213> Homo sapiens

<400> 88
Met Gly Glu Ala Ile Phe Arg Ser Ser Ser Cys Ser Gln Ser Thr Asn
1 5 10 15
Thr Ala Arg Val Ser Pro Arg Glu Thr Arg Arg Arg Thr Ile Arg Met
20 25 30
Arg Asn Lys Ser Leu Leu Leu Phe Ser Asp Thr Asn Met Lys Asn Glu
35 40 45
Gly Glu Asp Ile Thr Thr Gly Asn Ile Met His Asn Ser Met His Ile
50 55 60
Asn Ser Ala Thr Leu Met Lys Trp Thr Asn
65 70 74

<210> 89
<211> 92
<212> PRT
<213> Homo sapiens

<400> 89
Met Pro Leu Trp Tyr Phe Ser Asn Asp Thr Phe Lys Ser Met Val Ser
1 5 10 15
Phe Asp Cys Asn Arg Glu Asp Met Ala Arg Lys Met Trp Gln Asp His
20 25 30
Leu Leu Gly Glu His Ser Arg Ile Pro Ser Val Leu Val Pro Gln His
35 40 45
Gln Leu Thr Met Phe Leu Leu Met Met Thr Gly Arg Val Asp Thr Ser
50 55 60
Leu His His Ile Leu Phe Ser Asp Lys Val Gln Ala Thr Val Phe Ser

65 70 75 80
Thr Thr Glu Ser Ile Arg Phe Ser Ser Phe Ser Ser
85 90 92

<210> 90
<211> 240
<212> PRT
<213> Homo sapiens

<400> 90
Met Ala Ser Ser Asn Val Asn Asp Ile Met Pro Val Pro Glu Ala Leu
1 5 10 15
Val Tyr Val Gly Pro Arg Pro Thr Glu Asn Arg Cys Asn Val Trp Ser
20 25 30
Ile Ser Ser His Leu Gly Ser Val Asp Asp Leu Glu Asn Glu Leu Thr
35 40 45
Val Gln Leu Arg Asn Tyr Leu Glu Lys Ala Gly Asn Cys Ile Ser Thr
50 55 60
Trp Lys Arg Leu Arg Cys Ala Gly Pro Met Thr Val Lys Ala Glu Lys
65 70 75 80
Gly Ala Val Gln Pro Ala Glu Ala Arg Pro Gly Ser Arg Arg Gly Arg
85 90 95
Arg Thr Thr Ala Leu Val Ser Arg Leu Gly Ser Ser Val Ser His Arg
100 105 110
Val Leu Gly Leu Ala Val Ser Phe Ser Leu Gly Arg His Trp Pro Met
115 120 125
Ser Pro His Gly Asp Gly Arg Gly Gln Ala Gln Gly Arg Ala Val Arg
130 135 140
Val Gly Leu Arg Arg Ser Gly Thr Ile Arg Gly Gly Val Ala Val Phe
145 150 155 160
Ala Ala Val Ala Ala Val Phe Thr Phe Thr Leu Pro Pro Ser Val Pro
165 170 175
Gly Gly Asp Ser Gly Glu Leu Ile Thr Ala Ala His Glu Leu Gly Val
180 185 190
Ala His Pro Pro Gly Tyr Pro Leu Phe Thr Leu Val Ala Lys Leu Ala
195 200 205
Ile Thr Leu Phe Pro Phe Gly Ser Ile Ala Tyr Arg Val Asn Leu Leu
210 215 220
Cys Gly Leu Phe Gly Ala Val Ala Arg Leu Phe Phe Thr Val Phe Arg
225 230 235 240

<210> 91
<211> 73
<212> PRT
<213> Homo sapiens

<400> 91
Met Tyr Ile Ala Ile Glu Cys Arg Ile Trp Asp Thr Trp Ile Pro Phe
1 5 10 15
Leu Leu Pro Leu Cys Ser Leu Pro Ser Met Gly Arg Leu Ser Met Ile
20 25 30
Thr Thr Ser Gly Leu Trp Gln Arg Leu Asp Leu Pro Leu Gly Gly Gln
35 40 45
Thr Trp Gly Gln Ser Met Gln Leu His Gln Leu Ser His Leu Gln Asp
50 55 60
Leu Lys Gln Ala Pro Pro Ser Pro Met

65 70 73

<210> 92
<211> 484
<212> PRT
<213> Homo sapiens

<400> 92

```
Met Ala Ser Ile Ser Glu Pro Val Thr Phe Arg Glu Phe Cys Pro Leu
  1               5                  10                  15
Tyr Tyr Leu Leu Asn Ala Ile Pro Thr Lys Ile Gln Lys Gly Phe Arg
            20                  25                  30
Ser Ile Val Val Tyr Leu Thr Ala Leu Asp Thr Asn Gly Asp Tyr Ile
        35                  40                  45
Ala Val Gly Ser Ser Ile Gly Met Leu Tyr Leu Tyr Cys Arg His Leu
    50                  55                  60
Asn Gln Met Arg Lys Tyr Asn Phe Glu Glu Phe Asn Ser Leu Lys Ser
65                  70                  75                  80
Leu Lys Cys Phe Arg His Phe Ile Arg Lys Thr Arg Glu Ser Asn Asp
                85                  90                  95
Glu Pro Gln Gln Asp Leu Ser Lys Lys Ser Met His Arg Gly Thr Leu
            100                 105                 110
Val Thr Leu Leu Lys Ile Thr Ile Leu Lys Ser Val His Arg Gly Val
        115                 120                 125
Leu Val Thr Leu Leu Asn Ile Thr Ile Leu Lys Ser Met His Arg Gly
    130                 135                 140
Val Leu Val Thr Leu Leu Asn Ile Thr Ile Leu Lys Ser Met His Arg
145                 150                 155                 160
Gly Val Leu Val Thr Leu Leu Lys Ile Ala Val Leu Lys Ser Met His
                165                 170                 175
Arg Gly Val Leu Val Thr Leu Leu Lys Ile Thr Val Leu Lys Ser Met
            180                 185                 190
His Arg Gly Ile Leu Val Thr Leu Leu Lys Ile Thr Ile Leu Lys Ser
        195                 200                 205
Met His Arg Gly Val Leu Val Thr Leu Leu Lys Ile Thr Ile Leu Lys
    210                 215                 220
Ser Met His Arg Gly Val Leu Val Thr Leu Leu Lys Ile Thr Ile Leu
225                 230                 235                 240
Lys Ser Met His Arg Gly Thr Leu Val Thr Leu Leu Lys Ile Thr Val
                245                 250                 255
Leu Lys Ser Met His Arg Gly Thr Leu Val Thr Leu Leu Lys Ile Thr
            260                 265                 270
Ile Leu Lys Ser Met His Arg Gly Ile Leu Val Thr Leu Leu Lys Ile
        275                 280                 285
Thr Ile Leu Lys Ser Met His Arg Gly Ile Leu Val Thr Leu Leu Lys
    290                 295                 300
Ile Thr Ile Leu Lys Ser Met His Arg Gly Val Leu Val Thr Leu Leu
305                 310                 315                 320
Lys Ile Thr Ile Leu Lys Ser Met His Arg Gly Val Leu Val Thr Leu
                325                 330                 335
Leu Lys Ile Thr Val Leu Lys Ser Met His Arg Gly Val Leu Val Thr
            340                 345                 350
Leu Leu Lys Ile Thr Ile Leu Lys Ser Met His Arg Gly Thr Leu Val
        355                 360                 365
Thr Leu Leu Lys Ile Thr Ile Leu Lys Ser Met His Arg Gly Val Leu
    370                 375                 380
Val Thr Leu Leu Lys Ile Thr Ile Leu Lys Ser Val His Arg Gly Val
385                 390                 395                 400
Leu Val Thr Leu Leu Lys Ile Thr Val Leu Lys Ser Met His Arg Gly
                405                 410                 415
Thr Leu Val Thr Leu Leu Lys Ile Thr Ile Leu Lys Ser Met His Arg
```

```
            420                 425                 430
Gly Val Leu Val Thr Leu Leu Lys Ile Thr Val Leu Lys Ser Met His
        435                 440                 445
Arg Gly Thr Leu Val Thr Leu Leu Lys Ile Thr Ile Leu Lys Ser Met
    450                 455                 460
His Arg Gly Val Leu Val Thr Leu Leu Lys Ile Thr Ile Leu Lys Ile
465                 470                 475                 480
Ala Arg Glu Gln
            484
```

<210> 93
<211> 504
<212> PRT
<213> Homo sapiens

<400> 93

```
Met Ser Ser Thr Met Asn Gly Tyr Pro Gly Ile Gln Ser Asn Gly Leu
  1               5                  10                  15
Asn Ile Thr Tyr Gly Ala Phe Leu Asn Leu Lys Asn Leu Arg Glu Leu
             20                  25                  30
Leu Leu Glu Tyr Asn Gln Leu Pro Gln Ile Pro Ser Gly Leu Pro Glu
         35                  40                  45
Ser Leu Thr Glu Leu Met Leu Ile Gln Asn Asn Ile Tyr Asn Ile Thr
     50                  55                  60
Lys Glu Gly Ile Ser Arg Leu Ile Asn Leu Lys Asn Leu Tyr Leu Ala
 65                  70                  75                  80
Trp Asn Cys Tyr Phe Asn Lys Val Cys Glu Lys Thr Asn Ile Glu Asp
                 85                  90                  95
Gly Val Phe Glu Thr Leu Thr Asn Leu Glu Leu Leu Ser Leu Ser Phe
            100                 105                 110
Asn Ser Leu Ser His Val Pro Pro Lys Leu Ala Lys Ala Pro Cys Cys
        115                 120                 125
Lys Leu Phe Leu Ser Asn Thr Gln Ile Lys Tyr Ile Ser Glu Glu Asp
    130                 135                 140
Phe Lys Gly Leu Ile Asn Leu Thr Leu Leu Asp Leu Ser Gly Asn Cys
145                 150                 155                 160
Pro Arg Cys Phe Asn Ala Pro Phe Pro Cys Val Pro Cys Asp Gly Gly
                165                 170                 175
Ala Ser Ile Asn Ile Asp Arg Phe Ala Phe Gln Asn Leu Thr Gln Leu
            180                 185                 190
Arg Tyr Leu Asn Leu Ser Ser Thr Ser Leu Arg Lys Ile Asn Ala Ala
        195                 200                 205
Trp Phe Lys Asn Met Pro His Leu Lys Val Leu Asp Leu Glu Phe Asn
    210                 215                 220
Tyr Leu Val Gly Glu Ile Ala Ser Gly Ala Phe Leu Thr Met Leu Pro
225                 230                 235                 240
Arg Leu Glu Ile Leu Asp Leu Ser Phe Asn Tyr Ile Lys Gly Ser Tyr
                245                 250                 255
Pro Gln His Ile Asn Ile Ser Arg Asn Phe Ser Lys Leu Leu Ser Leu
            260                 265                 270
Arg Ala Leu His Leu Arg Gly Tyr Val Phe Gln Glu Leu Arg Glu Asp
        275                 280                 285
Asp Phe Gln Pro Leu Met Gln Leu Pro Asn Leu Ser Thr Ile Asn Leu
    290                 295                 300
Gly Ile Asn Phe Ile Lys Gln Ile Asp Phe Lys Leu Phe Gln Asn Phe
305                 310                 315                 320
Ser Asn Leu Glu Ile Ile Tyr Leu Ser Glu Asn Arg Ile Ser Pro Leu
                325                 330                 335
Val Lys Asp Thr Arg Gln Ser Tyr Ala Asn Ser Ser Ser Phe Gln Arg
            340                 345                 350
His Ile Arg Lys Arg Arg Ser Thr Asp Phe Glu Phe Asp Pro His Ser
```

```
        355                 360                 365
Asn Phe Tyr His Phe Thr Arg Pro Leu Ile Lys Pro Gln Cys Ala Ala
    370                 375                 380
Tyr Gly Lys Ala Leu Asp Leu Ser Leu Asn Ser Ile Phe Phe Ile Gly
385                 390                 395                 400
Pro Asn Gln Phe Glu Asn Leu Pro Asp Ile Ala Cys Leu Asn Leu Ser
                405                 410                 415
Ala Asn Ser Asn Ala Gln Val Leu Ser Gly Thr Glu Phe Ser Ala Ile
            420                 425                 430
Pro His Val Lys Tyr Leu Asp Leu Thr Asn Asn Arg Leu Asp Phe Asp
        435                 440                 445
Asn Ala Ser Ala Leu Thr Glu Leu Ser Asp Leu Glu Val Leu Asp Leu
    450                 455                 460
Ser Tyr Asn Ser His Tyr Phe Arg Ile Ala Gly Arg Asn Thr Ser Ser
465                 470                 475                 480
Arg Ile Tyr Ser Lys Phe His Lys Ser Lys Ser Phe Lys Leu Glu Pro
                485                 490                 495
Gln Gln His Leu Tyr Phe Asn Arg
            500                 504
```

<210> 94
<211> 583
<212> PRT
<213> Homo sapiens

<400> 94

```
Met Ala Asp Tyr Ile Ala Asn Cys Thr Val Lys Val Asp Gln Leu Gly
  1               5                  10                  15
Ser Asp Asp Ile His Asn Ala Leu Lys Gln Thr Pro Lys Val Leu Val
             20                  25                  30
Val Gln Ser Phe Asp Met Phe Lys Asp Lys Asp Leu Thr Gly Pro Met
         35                  40                  45
Asn Glu Asn His Gly Leu Asn Tyr Thr Pro Leu Leu Tyr Ser Arg Gly
     50                  55                  60
Asn Pro Gly Ile Met Ser Pro Leu Ala Lys Lys Lys Leu Leu Ser Gln
 65                  70                  75                  80
Val Ser Gly Ala Ser Leu Ser Ser Ser Tyr Pro Tyr Gly Ser Pro Pro
                 85                  90                  95
Pro Leu Ile Ser Lys Lys Lys Leu Ile Ala Arg Asp Asp Leu Cys Ser
            100                 105                 110
Ser Leu Ser Gln Thr His His Gly Gln Ser Thr Asp His Met Ala Val
        115                 120                 125
Ser Arg Pro Ser Val Ile Gln His Val Gln Ser Phe Arg Ser Lys Pro
    130                 135                 140
Ser Glu Glu Arg Lys Thr Ile Asn Asp Ile Phe Lys His Glu Lys Leu
145                 150                 155                 160
Ser Arg Ser Asp Pro His Arg Cys Ser Phe Ser Lys His His Leu Asn
                165                 170                 175
Pro Leu Ala Asp Ser Tyr Val Leu Lys Gln Glu Ile Gln Glu Gly Lys
            180                 185                 190
Asp Lys Leu Leu Glu Lys Arg Ala Leu Pro His Ser His Met Pro Ser
        195                 200                 205
Phe Leu Ala Asp Phe Tyr Ser Ser Pro His Leu His Ser Leu Tyr Arg
    210                 215                 220
His Thr Glu His His Leu His Asn Glu Gln Thr Ser Lys Tyr Pro Ser
225                 230                 235                 240
Arg Asp Met Tyr Arg Glu Ser Glu Asn Ser Ser Phe Pro Ser His Arg
                245                 250                 255
His Gln Glu Lys Leu His Val Asn Tyr Leu Thr Ser Leu His Leu Gln
            260                 265                 270
Asp Lys Lys Ser Ala Ala Ala Glu Ala Pro Thr Asp Asp Gln Pro Thr
```

```
        275                 280                 285
Asp Leu Ser Leu Pro Lys Asn Pro His Lys Pro Thr Gly Lys Val Leu
    290                 295                 300
Gly Leu Ala His Ser Thr Thr Gly Pro Gln Glu Ser Lys Gly Ile Ser
305                 310                 315                 320
Gln Phe Gln Val Leu Gly Ser Gln Ser Arg Asp Cys His Pro Lys Ala
                325                 330                 335
Cys Arg Val Ser Pro Met Thr Met Ser Gly Pro Lys Lys Tyr Pro Glu
            340                 345                 350
Ser Leu Ser Arg Ser Gly Lys Pro His His Val Arg Leu Glu Asn Phe
        355                 360                 365
Arg Lys Met Glu Gly Met Val His Pro Ile Leu His Arg Lys Met Ser
    370                 375                 380
Pro Gln Asn Ile Gly Ala Ala Arg Pro Ile Lys Arg Ser Leu Glu Asp
385                 390                 395                 400
Leu Asp Leu Val Ile Ala Gly Lys Lys Ala Arg Ala Val Ser Pro Leu
                405                 410                 415
Asp Pro Ser Lys Glu Val Ser Gly Lys Glu Lys Ala Ser Glu Gln Glu
            420                 425                 430
Ser Glu Gly Ser Lys Ala Ala His Gly Leu Tyr Ser Gly Ser Leu Cys
        435                 440                 445
Asn Ser Gly Leu Asn Ser Arg Leu Pro Ala Gly Tyr Ser His Ser Leu
    450                 455                 460
Gln Tyr Leu Lys Asn Gln Thr Val Leu Ser Pro Leu Met Gln Pro Leu
465                 470                 475                 480
Ala Phe His Ser Leu Val Met Gln Arg Gly Ile Phe Thr Ser Pro Thr
                485                 490                 495
Asn Ser Gln Gln Leu Tyr Arg His Leu Ala Ala Ala Thr Pro Val Gly
            500                 505                 510
Ser Ser Tyr Gly Asp Leu Leu His Asn Ser Ile Tyr Pro Leu Ala Ala
        515                 520                 525
Ile Asn Pro Gln Ala Ala Phe Pro Ser Ser Gln Arg Val Ile Arg Ala
    530                 535                 540
Pro Gln Tyr Lys Thr Val Gly Ser Ala Leu Pro Ser Ser Leu Lys Arg
545                 550                 555                 560
His Gly Gln Gln Ser Phe Thr Pro Tyr Pro Gly Val Leu Ala Tyr Arg
                565                 570                 575
Val Arg Ser Gln Tyr Phe Phe
            580             583
```

<210> 95
<211> 163
<212> PRT
<213> Homo sapiens

<400> 95

```
Met Glu Ser Ser Leu Ser Gly Lys Ser Ser Ala Gly Pro Ser Gly Ser
  1                 5                  10                  15
Asp Phe Phe Phe Ile Asn Val Glu Pro Lys Phe Ser Ala Arg Ala Val
             20                  25                  30
Trp Gly Phe Phe Cys Cys Thr Gln Val Arg His Gln Glu Val Ser Gln
         35                  40                  45
Val Asp Lys Ser Arg Ser Lys Gly Glu Ala Asn Ser His Arg Glu Glu
     50                  55                  60
Gly Lys Gly Thr Glu Ala Pro Arg Gly Arg Gly His Leu Asn Cys Trp
 65                  70                  75                  80
Gly Gln Lys Arg Pro Ile Gly Ser Cys Gln His Asp Pro Cys Ser Lys
                 85                  90                  95
Asp Lys Thr Ser Ile Leu Arg Ser Ala Leu Cys Pro Leu Glu Asn Thr
            100                 105                 110
Asn Thr Arg Ala Ala Pro Ala Ala Val Ile Pro Gln Gly Glu Glu Gly
```

```
        115                 120                 125
Pro Gln Gln Gly His Gly Thr Gln Ser Cys Gly Leu Gly Ala Cys Trp
    130                 135                 140
Ile Gln Pro Leu Pro Ser Ser Arg Cys Leu Pro Phe Gln Val Pro His
145                 150                 155                 160
Ser Ser Asn
        163
```

<210> 96
<211> 49
<212> PRT
<213> Homo sapiens

<400> 96

```
Met Pro Leu Ser Asp Pro Val Asn Ile Ala Lys Val Ser Lys Tyr Lys
  1               5                  10                  15
Gln Glu Tyr Pro Ile Arg Asp Asn Glu Gln Arg Leu Cys Ile Gln Ser
            20                  25                  30
Glu Lys Gly Thr Asn Arg Asp Lys Leu Thr Phe His Val Arg Asn Leu
        35                  40                  45
Gly
 49
```

<210> 97
<211> 111
<212> PRT
<213> Homo sapiens

<400> 97

```
Met Val Gln Lys Ala Pro Cys Trp Arg Val Ser Val Glu Leu Ala Leu
  1               5                  10                  15
Glu Leu Cys Gly Ser Pro Glu Phe Ser Met Phe Ala Asp Ala Ala Met
            20                  25                  30
Ala Ala Leu Ser Trp Ala Ala Pro Ser Ser Ala Ser Thr Cys Ser Val
        35                  40                  45
Pro Arg Ser Ser Ala Ala Gln Ala Pro Arg Ala Pro Ser Pro Ile Val
    50                  55                  60
Pro Arg Ala Gly Cys Gly Ser Leu Val Gly Ser Gln Pro Arg Trp Arg
65                  70                  75                  80
Arg Arg Glu Arg Leu Leu Cys Ala Ala Gly Ala Gly Ser Ala Ser Pro
                85                  90                  95
Ser Ala Ala His Val Ser Pro Pro Ala Ala Ser Cys Arg Ser Cys
            100                 105                 110 111
```

<210> 98
<211> 203
<212> PRT
<213> Homo sapiens

<400> 98

```
Met Glu Thr Trp Asp Ser Gly Leu Cys Phe Pro Gly Cys Leu Ser Leu
  1               5                  10                  15
Thr Gly Pro Gly Ser Val Thr Gly Thr Ala Gly Asp Ser Leu Thr Val
            20                  25                  30
Trp Cys Gln Tyr Glu Ser Met Tyr Lys Gly Tyr Asn Lys Tyr Trp Cys
```

```
         35                  40                  45
Arg Gly Gln Tyr Asp Thr Ser Cys Glu Ser Ile Val Glu Thr Lys Gly
     50                  55                  60
Glu Glu Lys Val Glu Arg Asn Gly Arg Val Ser Ile Arg Asp His Pro
 65                  70                  75                  80
Glu Ala Leu Ala Phe Thr Val Thr Met Gln Asn Leu Asn Glu Asp Asp
                 85                  90                  95
Ala Gly Ser Tyr Trp Cys Lys Ile Gln Thr Val Trp Val Leu Asp Ser
            100                 105                 110
Trp Ser Arg Asp Pro Ser Asp Leu Val Arg Val Tyr Val Ser Pro Ala
        115                 120                 125
Ile Thr Thr Pro Arg Arg Thr Thr His Pro Ala Thr Pro Pro Ile Phe
    130                 135                 140
Leu Val Val Asn Pro Gly Arg Asn Leu Ser Thr Gly Glu Val Leu Thr
145                 150                 155                 160
Gln Asn Ser Gly Phe Arg Leu Ser Ser Pro His Phe Leu Leu Val Val
                165                 170                 175
Leu Leu Lys Leu Pro Leu Leu Leu Ser Met Leu Gly Ala Val Phe Trp
            180                 185                 190
Val Asn Arg Pro Gln Trp Ala Pro Pro Gly Arg
        195                 200         203
```

<210> 99
<211> 65
<212> PRT
<213> Homo sapiens

<400> 99

```
Met Arg His Cys Pro Thr Phe Val Gln Ile Ser Ala Thr Val Phe Phe
  1                 5                  10                  15
Ser Phe Gly Pro Tyr Ile Lys Tyr Gly Ser Phe Gln Arg Lys Glu Thr
             20                  25                  30
Lys Val Ala Val Pro Pro His Ile Phe Arg Arg Arg Arg Gln Arg Leu
         35                  40                  45
Gly Ser Thr Arg Leu Leu Phe Trp Trp Val Trp Trp Gly Cys Ile Ile
     50                  55                  60
Thr
 65
```

<210> 100
<211> 174
<212> PRT
<213> Homo sapiens

<400> 100

```
Met Ser Arg Tyr Ser Tyr Gln Ser Leu Leu Asp Trp Leu Tyr Gly Gly
  1                 5                  10                  15
Val Asp Pro Ser Ser Ala Gly Asn Gly Gly Pro Asp Cys Ala Ala Phe
             20                  25                  30
Leu Ser Cys Gln His Arg Leu Leu Glu Ser Val Val Val Leu Thr Leu
         35                  40                  45
Ala Leu Leu Glu Ile Leu Val Ala Leu Arg His Ile Leu Arg Gln Thr
     50                  55                  60
Lys Glu Asp Gly Arg Gly Ser Pro Gly Ser Gln Pro Glu Gln Val Thr
 65                  70                  75                  80
Gln Arg Pro Glu Glu Gly Lys Glu Ser Leu Ser Lys Asn Leu Leu Leu
                 85                  90                  95
Val Ala Leu Cys Leu Thr Phe Gly Val Glu Val Gly Phe Lys Phe Ala
```

```
            100                 105                 110
Thr Lys Thr Val Ile Tyr Leu Leu Asn Pro Cys His Leu Val Thr Met
        115                 120                 125
Met His Ile Phe Leu Leu Ala Cys Pro Pro Cys Arg Gly Ala Ile Arg
    130                 135                 140
Arg Leu Gln Val Arg Ala Thr Asp Ala His Val Glu Trp Ser Ser Ser
145                 150                 155                 160
Gly Ile Ala Val Ser Cys Gly Lys His Ser Ala Ala Pro Leu
                165                 170             174
```

<210> 101
<211> 150
<212> PRT
<213> Homo sapiens

<400> 101

```
Met Glu Gly Gly Gln Val Pro Ser His Leu Pro Leu Ser His Ala Arg
  1                 5                  10                  15
Thr Cys Trp Ala Gly Leu Leu Gly Asn Thr Gly Asp Ser Asn Glu Leu
             20                  25                  30
Cys Ile Phe Asn Val Pro Ser Thr Ala Ser Gly Pro Gly Gly Pro Pro
         35                  40                  45
Asp Pro His Trp Leu Ala Pro Gln Pro Trp Ala Trp Pro His Leu Ser
     50                  55                  60
Trp Ser Pro Glu Pro Leu Ala Trp Ser Cys Leu Ile Cys Val Cys Val
 65                  70                  75                  80
Val Leu Ser Ala Pro Pro Leu Pro Leu Tyr Glu Phe Ser Cys Leu Ala
                 85                  90                  95
Gln Pro Leu Pro Cys Ile Ser Ala Ser Ile Arg Pro Leu Thr Glu Thr
            100                 105                 110
Thr Met Ser Cys Val Ser Leu Cys Gly Pro Leu Val Cys Leu Thr Gly
        115                 120                 125
His Gly Val Trp Asp Met Pro Gly Thr Arg Gly Ala His Trp Leu Arg
    130                 135                 140
His Leu Cys Pro Gly Ser
145                 150
```

<210> 102
<211> 100
<212> PRT
<213> Homo sapiens

<400> 102

```
Met Ala Leu Ala Leu Ala Ala Tyr Val Cys Gly Trp Val Val Asp Arg
  1                 5                  10                  15
Glu Thr Trp Pro Val Pro Met Pro Cys Asn Lys Gly Gly Arg Ala Cys
             20                  25                  30
Asn Leu Glu Met Gly Met Glu Trp Leu Asn Leu His Cys Glu Val Ser
         35                  40                  45
Lys Trp Gln Gln Pro Pro Ser Gly Ala Leu Cys Cys Ser Leu Ala Pro
     50                  55                  60
Leu Gln Ser Ile Phe Phe Pro Ala Ala Lys Val Met Phe Lys Asn Gly
 65                  70                  75                  80
Ser Trp Thr Val Leu Leu Pro Cys Ser Glu Phe Pro Ile Gly Phe Pro
                 85                  90                  95
Ser His Leu Glu
            100
```

<210> 103
<211> 143
<212> PRT
<213> Homo sapiens

<400> 103
Met Pro Glu Glu Asp Pro Ala Thr Leu Asp Asp His Ser Gly Thr Thr
1 5 10 15
Gly Thr Pro Lys His Phe Arg Asn Arg Lys Lys Ala His Phe Thr Ser
20 25 30
Phe Phe Ala Gly Asn Pro Arg Ile Gln His Phe Pro Gly Leu Pro Val
35 40 45
Arg Leu Arg Thr Gly Gly Pro Leu Cys Leu Gln Ala Phe Gly Gly Leu
50 55 60
Arg Ser Leu Gln Leu Leu Leu Gly Gly Leu Cys Gln Pro Leu Ser Gln
65 70 75 80
Gly Ile Gly Gln Arg Lys Gly Arg Gly Thr Leu Ala Leu Gly Phe Ser
85 90 95
Ser Gln Arg Ser Gly Arg Gly Asp Thr Gly Met Gly Leu Ala Leu Tyr
100 105 110
Arg Gln Pro Gly Ser Ser Pro Leu Asp Leu Arg Leu Leu Arg Thr Thr
115 120 125
Arg Glu Gly Asp Val Arg Thr Leu Ser Cys Pro Pro Asp Gly Pro
130 135 140 143

<210> 104
<211> 65
<212> PRT
<213> Homo sapiens

<400> 104
Met Asn Asn Ser Pro Leu Gly Ser Ser Phe Leu Gly Arg Leu Glu Glu
1 5 10 15
Val Pro Ser Pro Ser Pro Leu Pro Leu Val Ser Leu Pro Ser Gln Arg
20 25 30
Thr Ser Phe Ser Asn Ile Ala Pro Thr Thr Tyr Gln Tyr Cys Thr Pro
35 40 45
Gly Ser Cys Gln Ala Val His Ser Asn Ala Val Gly Gly Asn Thr Trp
50 55 60
Lys
65

<210> 105
<211> 56
<212> PRT
<213> Homo sapiens

<400> 105
Met Phe Tyr Leu Pro Ile Gln Lys Val Pro Pro Pro Ser Leu Pro Gln
1 5 10 15
Leu Thr Val Thr Ile Thr Leu Ala Ile Thr Ile Leu Val Ser Leu Pro
20 25 30
Pro Asp Gly Ser Gln Leu Leu Leu Ser Ser Trp Pro Pro Gln Trp Leu
35 40 45
Ser His Val Tyr Leu Pro Leu Cys

50 55 56

<210> 106
<211> 103
<212> PRT
<213> Homo sapiens

<400> 106

```
Met Phe Leu Thr Ile Leu Leu Asn Cys Val His Ser His His Ala Pro
 1               5                  10                  15
Val Phe Phe Gln Pro Arg Asn Ala Leu Thr Phe Asp Cys Gly Asp Asp
            20                  25                  30
Val Glu Val His Arg Val Phe Thr Asp Pro Val Ile Leu Phe Cys Asp
        35                  40                  45
Arg Gln Asn Gln Ser Ile Lys Arg Pro Ala Leu His Leu His Ile Pro
    50                  55                  60
Gly Thr Ser Pro His Phe Arg Ile Arg Asp His Ser Val Ile Gln Pro
65                  70                  75                  80
Asn Leu Ile Asn Lys Cys Ser Ser Pro Val Pro Gln Gln Ser Gln Met
                85                  90                  95
Lys Asp Val Ile Met Trp Asn
            100         103
```

<210> 107
<211> 65
<212> PRT
<213> Homo sapiens

<400> 107

```
Met Leu Pro Pro Leu Cys Trp Cys Cys Val Arg Thr Met Thr Cys Cys
 1               5                  10                  15
Ile Gly Thr Ser Thr Gly Met Asp Gly Arg Pro Pro Ser Pro Trp Arg
            20                  25                  30
Arg Ile Pro Cys Trp Thr Gln Thr Cys Ser Cys Arg Asn Ser Ala Thr
        35                  40                  45
Pro Ser Ser Pro His Phe Leu His Thr Ser Arg Trp Pro Gly Pro Met
    50                  55                  60
Tyr
65
```

<210> 108
<211> 109
<212> PRT
<213> Homo sapiens

<400> 108

```
Met His Ala Leu Phe Met Leu Phe Ser Leu Pro Ala Pro Gln Ser Leu
 1               5                  10                  15
Tyr Leu Gln Asn Gly Asp Asn Thr Pro Ser Ala Cys Leu Ile Gln Arg
            20                  25                  30
Arg Gly Ser Cys Leu Met Val Glu Phe Trp Ile Ser Asp Gln Val Arg
        35                  40                  45
Glu Gly Ser Gly Lys Val Thr Glu Gly Ser Asp Leu Gly Gln Glu Ala
    50                  55                  60
Ile Ser Gly Ala His Pro Ala Arg Asp Trp Gly Ser Leu Leu His Gly
```

```
 65                  70                  75                  80
Pro Cys Ile Pro Arg Gly Lys Arg Gly Pro Leu Trp Asn Pro Met Ala
                85                  90                  95
Pro Cys Ser Ile Ser Lys Cys Gln Gly Ser Trp Asn Val
            100                 105                 109
```

<210> 109
<211> 173
<212> PRT
<213> Homo sapiens

<400> 109

```
Met Lys Ala Pro Gln Ser Leu Val Lys Gly Ala Gly Asp Pro Glu Pro
  1               5                  10                  15
Asp Leu Trp Ile Ile Gln Pro Gln Glu Leu Val Leu Gly Thr Thr Gly
            20                  25                  30
Asp Thr Val Phe Leu Asn Cys Thr Val Leu Gly Asp Gly Pro Pro Gly
        35                  40                  45
Pro Ile Arg Trp Phe Gln Gly Ala Gly Leu Ser Arg Glu Ala Ile Tyr
    50                  55                  60
Asn Phe Gly Gly Ile Ser His Pro Lys Ala Thr Ala Val Gln Ala Ser
65                  70                  75                  80
Asn Asn Asp Phe Ser Ile Leu Leu Gln Asn Val Ser Ser Glu Asp Ala
                85                  90                  95
Gly Thr Tyr Tyr Cys Val Lys Phe Gln Arg Lys Pro Asn Arg Gln Tyr
            100                 105                 110
Leu Ser Gly Gln Gly Thr Ser Leu Lys Val Lys Gly Glu Thr Ile Gly
        115                 120                 125
Phe Gly Leu Val Gly Arg Ala Glu Gln Val Phe Asn Arg Asn Phe Ser
    130                 135                 140
Thr Cys Pro Glu Gln Ala Ser Ser Ser Gln Ala Ser Met Ser Leu His
145                 150                 155                 160
Lys Phe Phe Pro Leu Pro Gly Met Leu Ser Pro Phe Leu
                165                 170         173
```

<210> 110
<211> 64
<212> PRT
<213> Homo sapiens

<400> 110

```
Met Pro Asp Ser Gly Arg Leu His Trp Leu Leu Pro Ser Tyr Asp Gly
  1               5                  10                  15
Ala Ser Ser Val Val Leu Pro Ser Ser Gly Phe Pro Pro Leu Gly Asn
            20                  25                  30
Phe Leu Thr Ile Ser Thr Gly Arg Gln Arg Met Ala Ser Gly Pro Gly
        35                  40                  45
Thr Cys Val Phe Leu Ser Phe Met Phe Pro Asn Pro Thr Ile Gly Met
    50                  55                  60              64
```

<210> 111
<211> 42
<212> PRT
<213> Homo sapiens

<400> 111
Met Gln His Leu Arg Ser His Cys His Leu Phe Arg Gln Ala Leu Phe
1 5 10 15
Leu Cys Ser His Leu His Leu Ile Ser Trp Leu Ser Ser Tyr Leu Thr
20 25 30
Gly Val Asn Phe Leu Tyr Glu Leu Leu Phe
35 40 42

<210> 112
<211> 82
<212> PRT
<213> Homo sapiens

<400> 112
Met Lys Asn Glu Pro Lys Ser Thr Ile Ser Cys Pro Ser Ser Tyr Val
1 5 10 15
Leu Trp Gly Asn Val Val Arg His Lys Tyr His Leu His Met Gln Ala
20 25 30
Thr His Cys Lys Thr Leu Gly Gln His Ser Ser Glu Asn Leu Cys Cys
35 40 45
Leu Ser Thr Pro Gln Gly Ser His Cys Ile Val Ile Arg Pro Leu Pro
50 55 60
Trp Gly Gln Glu Ala Glu Val His Ser Cys Ser Val Met Asn Cys Leu
65 70 75 80
His Asp
82

<210> 113
<211> 79
<212> PRT
<213> Homo sapiens

<400> 113
Met Trp Ala Asp Ser Ile Leu Ala Ser Leu Leu Leu Trp Pro His Gln
1 5 10 15
Ser Leu Gln Leu Trp His His Pro His Leu Ala Asn Lys Asn Met Gly
20 25 30
Val Pro Pro Pro Thr Thr Cys Lys Pro Trp Ser Thr Glu Ala Gln Lys
35 40 45
Phe Ala Asp Tyr Ile Pro Phe Met Thr Thr Trp Pro Pro Leu Gly Lys
50 55 60
Arg Lys Val Val Lys Phe Asn Leu Ala Val His Arg Cys Ser Thr
65 70 75 79

<210> 114
<211> 169
<212> PRT
<213> Homo sapiens

<400> 114
Met Arg Leu Leu His Cys Lys Thr Leu His Ile Val Leu Phe Thr Leu
1 5 10 15
Leu Tyr Lys Ile Leu Met Asp His Gln Asn Leu Ser Glu His Val Leu

```
            20                  25                  30
Cys Met Val Leu Tyr Leu Ile Glu Leu Gly Leu Glu Asn Ser Ala Glu
        35                  40                  45
Glu Glu Ser Asp Glu Glu Ala Ser Val Gly Gly Pro Glu Arg Cys His
    50                  55                  60
Asp Ser Trp Phe Pro Gly Ser Asn Leu Val Ser Asn Met Arg His Phe
65                  70                  75                  80
Ile Asn Tyr Val Arg Val Arg Val Pro Glu Thr Ala Pro Glu Val Lys
                85                  90                  95
Arg Asp Ser Pro Ala Ser Thr Ser Ser Asp Asn Leu Gly Ser Leu Gln
            100                 105                 110
Asn Ser Gly Thr Ala Gln Val Phe Ser Leu Val Ala Glu Arg Arg Lys
        115                 120                 125
Lys Phe Gln Glu Ile Ile Asn Arg Ser Ser Ser Glu Ala Asn Gln Val
    130                 135                 140
Val Arg Pro Thr Thr Ser Ser Lys Trp Ser Ala Pro Gly Ser Ala Pro
145                 150                 155                 160
Gln Leu Thr Thr Ala Ile Phe Gly Asn
                165                 169
```

<210> 115
<211> 68
<212> PRT
<213> Homo sapiens

<400> 115

```
Met Lys Met Phe Gln Met Leu Leu Thr Ser Ser Phe Cys Ser Leu Ser
  1                 5                  10                  15
His Leu Gln Ser Cys Gln His Ile Ser Phe Leu Ser Ile Ser Asn His
            20                  25                  30
Ser Lys Ile Phe Leu Tyr Leu Gln Pro Thr Cys Tyr Leu Tyr Leu Pro
        35                  40                  45
Pro Leu Pro Leu Phe Ser Arg Ser Trp His Trp Asn Leu Arg Val His
    50                  55                  60
Ile Cys Ser Pro
65          68
```

<210> 116
<211> 415
<212> PRT
<213> Homo sapiens

<400> 116

```
Met Ser Met Lys Trp His Asn Trp Ile Trp Pro Trp Leu Phe Lys Thr
  1                 5                  10                  15
Lys Gly Cys Ser Leu Glu Ser Leu Lys Val Ala Thr Met Met Ser Lys
            20                  25                  30
Ala Thr Asp Lys Val Ser Lys Pro Arg Glu Gln Ile Ser Val Lys Ser
        35                  40                  45
Ser Val Pro Glu Ile Pro His Ala Asn Gly Leu Val Ser Val Asp Glu
    50                  55                  60
Glu Ser Tyr Gly Ala Leu Tyr Gln Ala Thr Val Glu Thr Thr Arg Ser
65                  70                  75                  80
Leu Ala Ala Ala Leu Asp Ile Leu Asn Thr Ala Thr Arg Asp Ile Gly
                85                  90                  95
Met Leu Asn Thr Ser Val Pro Asn Asp Met Asp Glu Gln Gln Asn Ala
            100                 105                 110
Arg Glu Ser Leu Glu Asp Gln Asn Leu Lys Asp Gln Asp His Leu Tyr
```

```
        115                 120                 125
Glu Glu Glu Ile Gly Ala Val Gly Gly Ile Asp Tyr Asn Asp Thr Asn
    130                 135                 140
Gln Asn Ala Gln Ser Glu Gln Asn Gly Ser Ser Asp Leu Leu Cys Asp
145                 150                 155                 160
Leu Asn Thr Ser Ser Tyr Asp Thr Ser Ala Leu Cys Asn Gly Phe Pro
                165                 170                 175
Leu Glu Asn Ile Cys Thr Gln Val Ile Asp Gln Asn Gln Asn Leu His
            180                 185                 190
Gly Asp Ser Lys Gln Ser Asn Leu Thr Asn Gly Asp Cys Val Ala Ser
        195                 200                 205
Ser Asp Gly Thr Ser Lys Pro Ser Ser Ser Leu Ala Val Ala Ala Gln
    210                 215                 220
Leu Arg Glu Ile Ile Pro Ser Ser Ala Leu Pro Asn Gly Thr Val Gln
225                 230                 235                 240
His Ile Leu Met Pro Asp Asp Glu Gly Glu Gly Glu Leu Cys Trp Lys
                245                 250                 255
Lys Val Asp Leu Gly Asp Val Lys Asn Val Asp Val Leu Ser Phe Ser
            260                 265                 270
His Ala Pro Ser Phe Asn Phe Leu Ser Asn Ser Cys Trp Ser Lys Pro
        275                 280                 285
Lys Glu Asp Lys Ala Val Asp Thr Ser Asp Leu Glu Val Ala Glu Asp
    290                 295                 300
Pro Met Gly Leu Gln Gly Ile Asp Leu Ile Thr Ala Ala Leu Leu Phe
305                 310                 315                 320
Cys Leu Gly Asp Ser Pro Gly Gly Arg Gly Ile Ser Asp Ser Arg Met
                325                 330                 335
Ala Asp Ile Tyr His Ile Asp Val Gly Thr Gln Thr Phe Ser Leu Pro
            340                 345                 350
Ser Ala Ile Leu Ala Thr Ser Thr Met Val Gly Glu Ile Ala Ser Ala
        355                 360                 365
Ser Ala Cys Asp His Ala Asn Pro Gln Leu Ser Asn Pro Ser Pro Phe
    370                 375                 380
Gln Thr Leu Gly Leu Asp Leu Val Leu Glu Cys Val Ala Arg Tyr Gln
385                 390                 395                 400
Pro Lys Gln Ala Phe Asn Val Tyr Leu Cys Val Trp Thr Val Ile
                405                 410                 415
```

<210> 117
<211> 77
<212> PRT
<213> Homo sapiens

<400> 117

```
Met Cys Thr Phe Arg Gly Leu Leu Thr Gly Leu Leu Thr Phe Pro Leu
 1               5                  10                  15
Phe Ser Pro Val Leu Tyr Phe Cys Asn Lys Phe Pro Asn Lys Thr Asn
            20                  25                  30
Met Phe Leu Leu Cys Phe Cys Lys Asn Tyr Phe Leu Ser Thr Val Tyr
        35                  40                  45
Lys Val Phe Val Arg Gln Gly Leu Val Phe Ala Ala His Thr Ala Glu
    50                  55                  60
Ala Met Val Ala Val Glu Val Ala Arg Lys Arg Thr Arg
65                  70                  75      77
```

<210> 118
<211> 55
<212> PRT
<213> Homo sapiens

<400> 118
Met Leu Gln Arg Gly Cys Asn Asp Trp Val Leu Pro Cys Val Ser Glu
1 5 10 15
Asn Gln Lys Gly Thr Ala Asp Leu Leu Val Phe Trp Ser Pro Val Val
20 25 30
Ser Ala Glu His Cys His Trp Cys Pro Asp Ser Arg Pro Lys Leu Gly
35 40 45
Lys His Cys Phe Trp Ser Cys
50 55

<210> 119
<211> 145
<212> PRT
<213> Homo sapiens

<400> 119
Met Ala Ser Asn Gly Leu Arg Gln Ser Thr Gly Pro Gly Arg Gly Pro
1 5 10 15
Ala Tyr Arg Gly Gln Gly Leu Glu Gln Ala Val Glu Ser Arg Val Pro
20 25 30
Glu Ala Ser Gln Ala Gly Glu Gly Gly Gln Glu Pro Gln Glu Glu Ala
35 40 45
Ala Val Asn Ile Leu Glu Tyr Leu Ala Leu Ile Gly Phe Trp Leu Phe
50 55 60
Val Glu Leu Val Ser Glu Val Leu Ser Thr Arg Ile Val Gly Ala Glu
65 70 75 80
Ala Glu Val Gly Val Glu Gly Pro Asp Pro Lys Glu Asp Ala Ser Arg
85 90 95
Asp Trp Ala Thr Ala His Thr Tyr Leu Cys Tyr Ile Lys Gln Glu Lys
100 105 110
Leu Gln Val Lys Gly Cys Gly Ala Gly Gly Tyr Leu Val Gly Asp Pro
115 120 125
Thr Asn Ala Ile Ala His Pro Gly Leu Pro Ala Glu Ala Ala Ile Ile
130 135 140
Tyr
145

<210> 120
<211> 74
<212> PRT
<213> Homo sapiens

<400> 120
Met Thr Ile Ser Gly Ser Gly Gln Glu Glu Ala Pro Arg Pro Ser Pro
1 5 10 15
Gln Ser Phe Gln Glu Asp Thr His Pro Phe Pro Val Leu Leu Phe Met
20 25 30
Leu Phe Leu Leu Ser Ser Lys Pro Ser Pro Leu Pro Phe Phe Leu Asn
35 40 45
Pro Thr Thr Gln Gly Pro Ala Leu Thr Thr Pro Gly Leu Ser Pro Phe
50 55 60
Ser Gly Glu Lys Thr Leu Arg Ser Leu Gln
65 70 74

<210> 121
<211> 77
<212> PRT
<213> Homo sapiens

<400> 121
Met Ser Ala Met Cys Glu Lys Cys Asp Lys Ser Leu Cys Met Lys Thr
1 5 10 15
Trp Leu Lys Gln Ser His Trp Gln Glu Ala Ala Glu Leu Pro Val Gln
20 25 30
Tyr Leu Val Arg Arg Met Arg Tyr Thr Ala Ile Leu Leu Gly Leu Phe
35 40 45
Gln Cys Gly Glu Gly Glu Glu Tyr Asn Arg Arg Pro Glu Arg Ala Lys
50 55 60
Arg Thr Pro Gln Val Thr Leu His Ile Arg Glu Gly Leu
65 70 75 77

<210> 122
<211> 71
<212> PRT
<213> Homo sapiens

<400> 122
Met Pro Ala Cys Cys Tyr Arg Pro Cys Leu Leu Gln Pro Ile Ser Leu
1 5 10 15
Leu Asn Ile Leu Leu Leu Leu Met Arg Lys Pro Ser Gln Glu Val Ile
20 25 30
Asn Asp Thr Pro Lys Ala Gly Lys Trp Leu Ser Arg Tyr Leu Asp Ser
35 40 45
Gly Leu Phe Tyr Ser Cys Ala Cys Gly Pro Phe Ser His Ser Phe Gln
50 55 60
Leu Asp Ser Phe Lys Asn Val
65 70 71

<210> 123
<211> 85
<212> PRT
<213> Homo sapiens

<400> 123
Met Arg Gln Ser Ser Cys Leu Gln Gly Pro Phe Leu Asp Leu Gly Arg
1 5 10 15
Ala Tyr Lys Phe Cys Val Leu Val Gly Ser Glu Gly Arg Gly Asp Gly
20 25 30
Ala Val Val Leu Ile Pro Ala Thr Gly His Ala Asn Ser Thr Ser Pro
35 40 45
Gly Ser Cys Pro Lys Gln Glu Gly Arg Arg Met His Gly Ile Gly Val
50 55 60
Gly Val Glu Met Arg Ala Ala Thr Trp Arg Glu Glu Asp Leu His Leu
65 70 75 80
Lys Ser Thr Gln Val
85

<210> 124
<211> 46

<212> PRT
<213> Homo sapiens

<400> 124
Met Ser Leu Gly Asp Asn Leu Leu Phe Lys Ala Lys Ile Thr Thr Val
1 5 10 15
His Cys Val Val Ser Val Lys Val Tyr Glu Asn Asn Ser Thr Lys Asn
20 25 30
Thr Glu Gly Arg Asn Ser Cys Gly Leu Glu Ala Val Lys Cys
35 40 45 46

<210> 125
<211> 61
<212> PRT
<213> Homo sapiens

<400> 125
Met Phe Met Pro Gly Thr Val Leu Arg Ile Leu Leu Ala Leu Pro Tyr
1 5 10 15
Leu Ile Leu Thr Lys Gln Val Gln Phe Phe Leu Phe Ser Asp Glu Ile
20 25 30
Met Ala Trp Lys Val Val Ala Pro Gly Leu Glu Leu Ser Ala Val Thr
35 40 45
Pro Asp Ser Thr Leu Phe Asn His Tyr Thr Ile Leu Ser
50 55 60 61

<210> 126
<211> 82
<212> PRT
<213> Homo sapiens

<400> 126
Met Gln Val Lys Phe Ile Leu Lys Tyr Tyr Ile Ser Phe Leu Trp Lys
1 5 10 15
Thr Val Thr Ala Asn Gly Glu Thr Val Asn Met Ser Leu Leu Tyr Ile
20 25 30
Phe Thr Thr Met Glu Met Arg Lys Lys Ser Glu Val Gly Leu His Leu
35 40 45
Pro Ile Ser Ile Leu Lys Pro Phe Phe Thr Ile Val Leu Asp Glu Lys
50 55 60
Ile Val Thr Gly Gln Val Trp Gly Gly Glu Leu Phe Leu Leu Phe Cys
65 70 75 80
Lys Asp
82

<210> 127
<211> 64
<212> PRT
<213> Homo sapiens

<400> 127
Met Val Leu Thr Ser Glu Asp Arg Gln Phe Tyr Leu Leu Thr Asp Ser
1 5 10 15
Phe Lys Leu Phe Tyr Gly Met Leu Lys Glu Thr Arg Ile Asn Lys Ser

```
             20                  25                  30
Thr Thr Phe Gly Leu Asp Cys Asn Lys Ile Phe His Lys Ser Lys Asp
         35                  40                  45
Leu Leu Gly Asn Glu Gly Lys Gln Arg Gly Cys Lys Asn His Leu Lys
     50                  55                  60                  64
```

<210> 128
<211> 81
<212> PRT
<213> Homo sapiens

<400> 128

```
Met Arg Cys Val Arg Arg Met Cys Met Trp Ser Ala Gly Pro Ala Pro
  1               5                  10                  15
Ala Thr Val Cys Ala Val Met Val Ala Gly Pro Lys Ser Pro Gln Phe
             20                  25                  30
Pro Pro Arg Trp Gly Cys Val Tyr Asn Leu Ile Gly Cys His Phe Leu
         35                  40                  45
Asp Pro Phe Ser Gly Tyr Phe Ser Gly Ile Ser Trp Arg Asp Ile Ser
     50                  55                  60
Phe Ser Leu Tyr Phe Met Gly Gln Glu Phe Gln Asn Gln Ser Leu Leu
 65                  70                  75                  80
Lys
 81
```

<210> 129
<211> 73
<212> PRT
<213> Homo sapiens

<400> 129

```
Met Cys Pro Leu Leu Val Tyr Lys Ile Ile Leu Val Phe Ala Ala Met
  1               5                  10                  15
Phe Phe Phe Ser Gln Gly Ser Gln Val Glu Ile Arg Ser His Glu Gly
             20                  25                  30
Glu His Cys Val Gly Thr Val His Leu Leu Ser His Phe Leu Tyr Ser
         35                  40                  45
Lys Asn Asn Pro Val Phe Tyr Lys Gly Asn Thr Ser Phe Ile Phe Glu
     50                  55                  60
Thr Met Glu Glu Asp Ser Leu Ser Ser
 65                  70                  73
```

<210> 130
<211> 60
<212> PRT
<213> Homo sapiens

<400> 130

```
Met Ala Val Ser Glu Leu Asp Arg His Thr Phe Asn Arg Gln Gly Cys
  1               5                  10                  15
Gly Lys Ser Lys Ile Lys Ala Gly Lys Arg Thr Gln His Val Ser Ser
             20                  25                  30
Leu Ser Ser Leu Ser Ala Lys Trp Leu Arg Val Glu Met Met Arg Glu
```

```
         35                  40                  45
Gln Glu His Lys Glu Asp Pro Leu Lys Gln Ser Gln
     50                  55                  60
```

<210> 131
<211> 101
<212> PRT
<213> Homo sapiens

<400> 131

```
Met Lys Lys Lys Pro Gln Leu Arg Arg Val Val Pro His Gly His Val
  1               5                  10                  15
Ala Val Gly Glu Ser Thr Leu Ile Tyr Arg Val Glu Glu Lys Arg Ala
            20                  25                  30
Pro Gly Ile His Gln Ala Thr Ala Leu Gln Lys Glu Ser His Ala Gln
         35                  40                  45
Glu Lys Gln Ile Ser Asp Ser Ala Thr Arg Lys Gly His Glu Ala Leu
     50                  55                  60
Lys Lys Arg Pro Met Pro Glu Ala Phe Lys Phe Lys Val Asp Cys Ala
 65                  70                  75                  80
Lys Phe Ser Met Ile Val Thr Ser Ile Gln Arg Lys Lys Lys Pro Leu
                85                  90                  95
Pro Thr Glu Leu Thr
            100 101
```

<210> 132
<211> 46
<212> PRT
<213> Homo sapiens

<400> 132

```
Met Leu Ala Thr Leu Ser Ile Gly Thr Leu Ser Met Leu Ile Ile Val
  1               5                  10                  15
Val Ser Asp Ser Trp Pro Tyr Ser Ser Asn Ser Pro Ala Met Phe Gly
            20                  25                  30
Ser Asp Ala Gly Phe Ile Pro Ser Asn Cys Ile Phe Ala Phe
         35                  40                  45  46
```

<210> 133
<211> 76
<212> PRT
<213> Homo sapiens

<400> 133

```
Met Arg Phe Ile Ser Pro Lys Ala Pro Cys Leu His Phe Ser Thr Asp
  1               5                  10                  15
His Thr Pro Val Gly Pro Cys Ala Trp Leu His Ala Ile Ser Leu Leu
            20                  25                  30
Arg Val Ile Arg Val Ala Cys Ser Ser Gly Leu Phe Phe Leu Cys Cys
         35                  40                  45
Ala Ser Val Asn Arg Arg Asn Leu Pro Pro Phe Met Pro Met Leu Trp
     50                  55                  60
Gly Cys Ser Pro Pro Gly Ala Val Ile Thr Gln Pro
 65                  70                  75  76
```

<210> 134
<211> 53
<212> PRT
<213> Homo sapiens

<400> 134
Met Ile Leu Val Ser Leu Leu Ile Leu Ile Val Glu Pro Leu Phe Ala
1 5 10 15
Ser Leu Thr Pro Leu Ser Leu Cys Phe Glu Cys Val Val Phe Leu Asn
20 25 30
Val Gly Gln His Leu Thr Asp Gln Thr Phe Ser Leu Asn Gly Leu Leu
35 40 45
Phe Leu Ser Asn Ser
50 53

<210> 135
<211> 67
<212> PRT
<213> Homo sapiens

<400> 135
Met Ser Pro Leu Leu Pro Leu Ser Tyr Lys Leu Val Leu Cys Phe Pro
1 5 10 15
Thr Pro Asn Gly Val Val Thr His Gly Glu Gln Asn Ala Ser Ser Thr
20 25 30
Asp Ile Glu His Gly Leu Lys Thr Ile Leu Ile Lys Pro Pro Ala Arg
35 40 45
Ile Leu Lys Arg Lys Thr Glu Gly Glu Glu Ser Asn Arg Leu Thr Leu
50 55 60
Pro Thr Thr
65 67

<210> 136
<211> 86
<212> PRT
<213> Homo sapiens

<400> 136
Met Gly Pro Val Ser Gly Cys Trp His Met Ser Leu Cys Leu Arg Val
1 5 10 15
Tyr Leu Ala Leu Asp Pro Ala His Gln Glu Leu Met Pro Pro Gly Ser
20 25 30
Thr Leu Gln Pro Ile Thr Leu Arg Ile Gly Ile Glu Ile Leu Gln Pro
35 40 45
Pro Thr Leu Glu Val Gly Asn Ser Glu Ala Leu Ser Val Pro Ser Arg
50 55 60
Thr Thr Pro Arg Arg Thr Glu Leu Pro Trp Pro Thr Glu Leu Thr Gly
65 70 75 80
Phe Leu Ile Asn Thr Leu
85 86

<210> 137
<211> 52

<212> PRT
<213> Homo sapiens

<400> 137

```
Met His Thr Cys Val Pro Glu Arg Leu Phe Gln Cys His His Leu Ile
  1               5                  10                  15
Arg Ile Thr Cys Leu Phe Met Ile Leu Glu Phe Arg Leu Phe Lys Tyr
             20                  25                  30
Gly Ser Asn Val Cys Ser His Val Ile Ile Asn His Pro Gln Val Gln
         35                  40                  45
Gly Arg His Arg
     50      52
```

<210> 138
<211> 114
<212> PRT
<213> Homo sapiens

<400> 138

```
Met Gly Gln Val Thr Cys Ser Gln Val Asn Ala Ser His Thr Tyr Asp
  1               5                  10                  15
Leu Thr Ser Ala Pro His Met Asn Asn Ser Ala Gln Pro Tyr Ala His
             20                  25                  30
Ser Gln Gly Ser Thr Cys Val Ala Ser Phe Ser Arg Leu Met Ser Gly
         35                  40                  45
Ile Pro Leu Ser Asn Ile Ser Ala Ile Phe Tyr Glu Thr Ile Phe Leu
     50                  55                  60
Asp Gly Gln Phe Leu Ala Ser Gln Glu Thr Asp Glu Glu Ala Arg Arg
 65                  70                  75                  80
Met Tyr Trp Glu Ser Pro Ala Glu Glu Glu Leu Leu Arg Ser Glu Trp
                 85                  90                  95
Arg Ser Tyr Leu Leu His Val Gly Arg Ser Cys Ser Ala Val Tyr Ser
            100                 105                 110
Ile Phe
    114
```

<210> 139
<211> 60
<212> PRT
<213> Homo sapiens

<400> 139

```
Met Ser Leu Arg Ile Arg Ala Ala Arg Asn Trp Ala Arg Asp Val Gln
  1               5                  10                  15
Lys Leu Trp Thr Ile Val Val Leu Leu Val Leu Ile Leu Ile Arg Ser
             20                  25                  30
Ala Val Asn Leu Leu Ile Asn Ser Arg Thr Glu Asp Lys Ser Leu Gln
         35                  40                  45
Leu Val Leu Tyr Gln Ser Val Ile Ile Cys Phe Pro
     50                  55                  60
```

<210> 140
<211> 58
<212> PRT
<213> Homo sapiens

<400> 140
Met Val Trp His Val Arg Lys Ser Ser Phe Val Trp Leu Leu Gln Leu
1 5 10 15
Phe Ser Phe Ile Ser Cys His Ser Val Ile Ser Val Ser Pro Val His
20 25 30
Val Pro Trp Thr Gln Cys Ala Val Ile Pro Pro Tyr Thr Ser Cys Pro
35 40 45
Lys Leu Phe Ala Ile Gln Gly Gly Arg Phe
50 55 58

<210> 141
<211> 80
<212> PRT
<213> Homo sapiens

<400> 141
Met Gly Gly Phe Lys Val Asn Leu Phe Val Lys Val Lys Ala Glu Gly
1 5 10 15
Ser Pro Leu Cys Trp Leu Lys Leu Ala Cys Leu Gly Ala Trp Leu Leu
20 25 30
Ser Leu Leu Ile Ser Gln Lys Ser Asp Glu Gln Leu Cys Phe Gly Leu
35 40 45
Leu Thr Trp Asn Phe Ser Thr Ser Asp Ser Ile Leu Val Trp Phe Val
50 55 60
Gly Pro Arg Ala Gly Thr Gln Ser Lys Pro Val Val Ser Tyr Lys Phe
65 70 75 80

<210> 142
<211> 58
<212> PRT
<213> Homo sapiens

<400> 142
Met Phe Ser Cys Ile Thr Thr Phe Ser Thr Glu His Asn Ile Tyr Ser
1 5 10 15
Ser Lys Lys Glu Lys Arg Lys Leu Leu Gln Thr Pro Asp Pro Asp Gln
20 25 30
Phe Pro Arg Asp Cys Thr Ile Thr Leu Gln Ala Ala Leu Thr Leu Pro
35 40 45
Pro Ala Ile Lys Lys Gln Ile Asn Pro Ile
50 55 58

<210> 143
<211> 72
<212> PRT
<213> Homo sapiens

<400> 143
Met Cys Ala Phe Leu Leu Pro Phe Lys Leu Leu Phe Phe Leu Glu Ile
1 5 10 15
Ser Leu Ala Met Lys Ser His Phe Pro Phe Thr Leu Leu Met Leu Ser

```
             20                  25                  30
Arg Val Leu Leu Lys Lys Thr Val Asp Gly Leu Lys Leu Gly Trp Val
         35                  40                  45
Ile Thr Ile Pro Ser Asn Glu Leu Thr Ser Val Phe Thr Leu Met Ile
     50                  55                  60
Asn Arg Pro Asn Gln Lys His Phe
 65                  70      72
```

<210> 144
<211> 69
<212> PRT
<213> Homo sapiens

<400> 144

```
Met Pro Thr Leu Leu Gln Val Met Ser Trp Met Leu Ser Phe Gly Thr
  1               5                  10                  15
Gln Thr Leu Gln Leu Glu Ser Cys Thr Cys Ala Leu His Ile Val Gly
             20                  25                  30
Ala Trp Lys Val Pro Tyr Pro Leu Phe Ser Arg Val Leu Ile Cys Gln
         35                  40                  45
Val Lys Ile Leu Ser Thr Ser Ile Ser Gln Glu Lys Val Phe Arg Thr
     50                  55                  60
Glu Ser Arg Thr Glu
 65              69
```

<210> 145
<211> 45
<212> PRT
<213> Homo sapiens

<400> 145

```
Met Cys Arg Ala Leu Leu Leu Leu Leu Cys Ser Pro Asn Ser Ser Phe
  1               5                  10                  15
Gln Trp Leu Pro Leu Pro Val His Pro His Thr Thr Ile Arg Tyr Arg
             20                  25                  30
Ser Tyr Asn Met Val Pro Val Lys Leu Thr Asn Val Gln
         35                  40                  45
```

<210> 146
<211> 56
<212> PRT
<213> Homo sapiens

<400> 146

```
Met Ile Ser Met Leu Pro Arg Pro Ser Val Thr Asn Thr Asn Cys Phe
  1               5                  10                  15
Tyr Leu Ala Leu Gln Thr Cys Phe Leu Arg Gly Arg Ala Arg Leu Cys
             20                  25                  30
Lys Val Lys Lys Asn Arg Thr Lys Pro Asn Thr Arg Thr Leu Arg Leu
         35                  40                  45
Phe Tyr Thr Pro Leu Phe Ser Gly
     50                  55  56
```

<210> 147
<211> 44
<212> PRT
<213> Homo sapiens

<400> 147
Met Tyr His Asn Val Ser Thr Gly Leu Leu Asn Ile Arg His Gln Ser
1 5 10 15
Gln Asn Ser Leu Ile Gly Ser Thr Cys Thr Cys Leu Val Asn Ala Ser
20 25 30
Met Asp Met Ser Ile Gln Leu Val Thr Tyr Arg Thr
35 40 44

<210> 148
<211> 60
<212> PRT
<213> Homo sapiens

<400> 148
Met Ile Gly Phe Ile Ile Phe Pro Ser Glu Ile Met Ala Thr Arg Ser
1 5 10 15
Lys Gly Ala Phe Ile Asn Cys Tyr Ile Ile Leu Leu Leu Thr Phe Leu
20 25 30
Met Ile Arg Thr Phe Tyr Asn Leu Met Glu Tyr Tyr Cys Pro Thr Leu
35 40 45
Leu Ile Arg Lys Leu Met Ser Asn Thr Lys Ile Leu
50 55 60

<210> 149
<211> 100
<212> PRT
<213> Homo sapiens

<400> 149
Met Lys Lys Leu Glu Phe Ser Met Arg Leu Gly Gly Trp Val Val Gly
1 5 10 15
Met Glu Leu Ser Gln Ser His Lys Val Asn Gly Pro Ser Ser Gln Phe
20 25 30
Leu Gly Ser Val Val Gln Leu Met Thr Arg Ser Trp Ala Trp Leu Phe
35 40 45
Glu Arg Val Ile Ser Cys Gly Ala Gln Val Lys Thr Thr Gly Thr Gly
50 55 60
Glu Phe Lys Val Asn Met Trp Leu Gly His Leu Ser Val Pro Gln Ser
65 70 75 80
Ser His Asp Ala Val Ser Val Ala Gly Lys Ser Tyr Arg Gly Leu Leu
85 90 95
Glu Asn Val Glu
100

<210> 150
<211> 85
<212> PRT
<213> Homo sapiens

<400> 150

```
Met Asp Phe Leu Ser Arg Leu Met Leu Leu Arg Met Cys Lys Cys Val
 1               5                  10                  15
Thr Ala Thr Tyr Gln Tyr Ile Arg Arg Ser Leu Phe Leu Asn Leu Val
            20                  25                  30
Pro Leu Leu Gln Thr Leu Ser Ile Leu Thr Ala His Ser Val Leu Leu
        35                  40                  45
Arg Pro Ala Leu Ser Ser Leu Val Lys Met Glu Asp Ser Gln Ala Leu
    50                  55                  60
Ser Leu Ser Leu Glu Pro Val Ser Ala Ile Leu Thr Val Pro Pro Glu
65                  70                  75                  80
Asp Ser Lys Val Leu
                85
```

<210> 151
<211> 133
<212> PRT
<213> Homo sapiens

<400> 151

```
Met Val Gln Ala Arg Gly Phe Leu Leu Pro Gln Asn Arg Leu Ser Gly
 1               5                  10                  15
Gln Gly Cys Leu Cys Ala Cys Asp Ser His Glu Thr Asn Ser Leu Cys
            20                  25                  30
Ala Val Thr Thr His Trp Phe Gln Arg Gln Asn Ile His Gly Lys Ser
        35                  40                  45
Arg Leu Lys Ser Arg Glu Gly Arg Val Ala Arg Arg Ile Arg Pro Gly
    50                  55                  60
Leu Asp Arg Pro Gly Gly Asn Thr Asp Pro Ser Arg Phe Trp Leu Pro
65                  70                  75                  80
Gly Ala Leu Ser His Ile Gly Leu Asp Leu Asn Arg Leu Tyr Ser Leu
                85                  90                  95
Ile Phe Pro Ser Pro Pro Leu Ser Leu Asp Pro Leu Pro Pro His Pro
            100                 105                 110
Thr Ser Leu Cys Ser Cys His Phe Leu Gln Ala Lys Ser Gly Glu Ile
        115                 120                 125
Val Asp Pro Gly Ile
    130             133
```

<210> 152
<211> 57
<212> PRT
<213> Homo sapiens

<400> 152

```
Met Arg Met Thr Thr Ser Ser Leu Val Leu Pro Pro Leu Phe Val Leu
 1               5                  10                  15
Lys Cys Gln Arg Phe Tyr Pro Pro Leu Tyr Leu His Pro Tyr Ser Ile
            20                  25                  30
Cys Gln His Val Ser Ile Leu Val Lys Asn Ser Leu Gly Pro Gly Gly
        35                  40                  45
Ser Lys Val Pro Thr Leu Gly Asn His
    50                  55      57
```

<210> 153
<211> 126

<212> PRT
<213> Homo sapiens

<400> 153

```
Met Gly Lys Gln Leu Ser Cys Tyr Leu Leu Gly Ile Trp Gln Val Pro
 1               5                  10                  15
Val Pro Cys Leu His Cys Leu Pro His Gly His Leu Ser Val Leu Glu
            20                  25                  30
Pro Leu Gly Gly Cys Val Leu Lys Arg Ile Leu Arg Gly Leu Ala Val
        35                  40                  45
Arg Thr Leu Ser Leu Leu Ser Val Leu Ser Pro Asp Arg Ala Ala Val
    50                  55                  60
Val Gly Ser Pro Arg Pro His Gly His Lys Cys Leu Leu Ala Ser Ile
65                  70                  75                  80
Arg Val Asp Val Pro Asp Ser Cys Ser Gly Leu Arg Leu Pro Gly Pro
                85                  90                  95
Pro His Ser Ala Ser Phe Leu His Ile Leu Gly Thr Asn Leu Cys Ile
            100                 105                 110
Ser Cys Leu Ala Pro Pro Gln Asn Val Leu Thr Glu Gly Arg
        115                 120                 125 126
```

<210> 154
<211> 91
<212> PRT
<213> Homo sapiens

<400> 154

```
Met Val Phe Val Gly Lys Cys Thr Gln Ser Ser Asn Leu Ser Phe Leu
 1               5                  10                  15
Arg Phe Leu Ile Pro Val Thr Met Ala Lys Gly Thr Pro Pro Ser Leu
            20                  25                  30
Phe Phe His Leu Ser Pro Gly Gly Lys Ile Lys Val His Arg Asp Thr
        35                  40                  45
Phe Pro Phe Thr Phe Pro His Ala Asn Arg Asp Pro Thr Leu Asp Cys
    50                  55                  60
Arg Ala Met Gly Arg Gly Asp Pro Gly Ser Ser Trp Leu Cys Asp Asn
65                  70                  75                  80
Arg Gly Lys Arg Arg Gly Gly Arg Glu His Arg
                85                  90  91
```

<210> 155
<211> 96
<212> PRT
<213> Homo sapiens

<400> 155

```
Met Leu Pro Ser Phe Leu Pro Gln Ser Leu Gly Asn Leu Ile His Thr
 1               5                  10                  15
Leu Gly Phe Leu Leu Ile Ile His Lys Tyr Met Ser Ala Phe Lys Asn
            20                  25                  30
Arg Thr Asp Glu Phe Met Asn Met Gly Met Gln Pro Tyr Ile Lys Ser
        35                  40                  45
Pro Tyr Arg Leu Ser Met Ser Gln Ile Ser Leu Lys Phe Asp Leu Ser
    50                  55                  60
Gln Thr Asp Leu Ile Leu Pro His Lys Phe Tyr Ser Pro Ser Ser Phe
65                  70                  75                  80
Pro Thr Val Met Leu Phe Tyr Ser Phe Gly Arg Leu Ser His Lys Pro
```

85 90 95 96

<210> 156
<211> 63
<212> PRT
<213> Homo sapiens

<400> 156
Met Gly Tyr Thr Val Lys Arg Glu Cys Leu Leu Ile Val Leu Met Gln
1 5 10 15
Ala Trp Arg Ser Phe Val Met Gly Val Glu Val Leu Met Tyr Ile Val
20 25 30
Ala Val Arg Cys Arg Ala Asp Phe Ala Thr Ser Leu Trp Gln Pro Trp
35 40 45
Cys Tyr Thr Arg Ala Gly Gly Gln Phe Asn Val Ser Gln Ala Arg
50 55 60 63

<210> 157
<211> 77
<212> PRT
<213> Homo sapiens

<400> 157
Met Val Val Leu Tyr Ile Val Arg Ala Tyr Asn His Tyr Ile Leu Cys
1 5 10 15
Cys Leu Ser Ser Ser Leu Tyr Leu Val Leu Ile Leu Leu Val Thr Val
20 25 30
Tyr Leu Met Leu Thr Thr Ser Ser Tyr Asn Asp Val Ser Leu Val Ile
35 40 45
Trp Ile Ala Ser Ser Phe Ala Ser Ser Lys Phe Phe Arg Lys Gly Leu
50 55 60
Arg Glu Tyr Ser Tyr Phe Met Asn Phe Leu Ala Arg Ser
65 70 75 77

<210> 158
<211> 92
<212> PRT
<213> Homo sapiens

<400> 158
Met Leu Glu Trp Pro Leu Leu Gly Gln Ile Leu Pro Met Ile Ile Pro
1 5 10 15
Leu Pro Pro Leu Pro Ala Leu Val Val Trp Pro Ile Gly Leu Thr His
20 25 30
Cys Pro Trp Pro Ser Pro Phe Val Pro Ala Ser Leu Asp Gly Phe Tyr
35 40 45
Asn Ser Arg Ser Leu Asp Gly Lys Ser Pro Pro Leu Arg Pro Glu Lys
50 55 60
Trp Ser Pro Trp Ser Trp Phe Ser Arg Leu Pro Gly His Gly Leu Pro
65 70 75 80
Lys Glu Gly Gly Gln Arg Lys Arg Glu Val Trp Ala
85 90 92

<210> 159
<211> 56
<212> PRT
<213> Homo sapiens

<400> 159
Met Met Leu Trp Gln Val Tyr Pro Gly Pro Ser Ala Gln Glu Leu Cys
1 5 10 15
Leu Phe Leu His Ala Pro Trp Ser Met Ser Thr Ala Val Glu Arg Asp
20 25 30
Lys Arg Gln Lys Asp Gly Arg Ser His Arg Met Leu Leu Pro Gln Pro
35 40 45
Gln Ser Leu Met Ser Arg Cys Cys
50 55 56

<210> 160
<211> 66
<212> PRT
<213> Homo sapiens

<400> 160
Met Ile Glu Lys Lys Ser Met Phe Phe Cys Leu Thr Arg Gln Ser Leu
1 5 10 15
Leu Cys Thr Leu Leu Leu Met Leu Lys Arg Cys Ile Phe Phe Ser Tyr
20 25 30
Cys Val Ile Cys Arg Ala Lys Ser Phe Glu Leu Phe Thr Ser Glu Ile
35 40 45
Thr Phe Pro Asp Lys Arg Ala Lys Gln Cys Leu Phe Lys Leu Phe Ser
50 55 60
Gly Thr
65 66

<210> 161
<211> 109
<212> PRT
<213> Homo sapiens

<400> 161
Met His Pro Tyr Ser Pro Ser Leu Leu Val Pro Gly Arg Gly Asn Thr
1 5 10 15
His Leu Ala Arg Ser Val Pro Pro His Thr Leu Pro Ser Pro Pro His
20 25 30
Ser Arg Tyr Pro Ala Leu Pro Thr Leu His Arg Trp Gly Pro Ala Pro
35 40 45
Arg Pro Pro Pro Leu Pro Val Ser Ala Ala Ala Pro Pro Gly His His
50 55 60
Thr Pro Ser Arg Pro Pro Pro Pro Ser Ser Leu Cys Ala Lys Pro His
65 70 75 80
Ile Pro Pro Leu Pro Arg Pro Arg Ile Thr Ala Arg Ser Pro Thr His
85 90 95
Pro Ala Arg Pro Ser Ala Ser Pro Arg Ala Pro His Arg
100 105 109

<210> 162
<211> 69
<212> PRT
<213> Homo sapiens

<400> 162
Met Ser Ile Leu Val Val Ser Ala Phe Leu Ala Asn Leu Arg Leu Leu
1 5 10 15
Met Thr Ile Ser Thr Ser Gln Met Leu Asn Met Thr Lys Ile Thr Tyr
20 25 30
Leu Val Leu Phe Leu His Leu Ser Ala Leu Arg Ile Gly Ser Thr Pro
35 40 45
His Ser Phe Leu Leu Lys Ser Tyr His Leu Gly Thr His Phe Ser Leu
50 55 60
Phe His Met Asn Ser
65 69

<210> 163
<211> 43
<212> PRT
<213> Homo sapiens

<400> 163
Met Phe Val Phe Val Ser Ile His Thr Glu Leu Val Pro Ile Leu Arg
1 5 10 15
Pro Leu Cys Leu Leu Tyr Cys Cys Pro Asp Cys Ser Val Pro Arg Pro
20 25 30
Leu Tyr Ser Leu Lys Tyr Leu Leu Leu Glu Lys
35 40 43

<210> 164
<211> 52
<212> PRT
<213> Homo sapiens

<400> 164
Met Met Arg Val Ile Ile Leu Ile Trp Phe Arg Ile Ser Lys Gly Thr
1 5 10 15
Phe Gln His Ser Thr Thr Lys Cys Asp Val Cys Phe Arg Val Phe Leu
20 25 30
Leu Ser Asn Cys Ser Phe Leu Ser Leu Asn Tyr Lys Leu Thr Ser Asp
35 40 45
Phe Ile Ile Tyr
50 52

<210> 165
<211> 62
<212> PRT
<213> Homo sapiens

<400> 165
Met Gln Met Val Val Pro Arg Leu Leu Ser Val Pro Gln Leu Leu Asn
1 5 10 15
Thr Ala Pro Leu Phe Leu Pro Trp Glu Lys Thr Val Lys Thr Gln Tyr

```
            20                  25                  30
Ser Gly Ile Ile Phe Lys Phe Lys Ser Arg Ile Glu Thr Ala Glu Lys
        35                  40                  45
Ser Ile Gly Asp Thr Lys Glu Arg Ile Gln Pro Ser Gln Ile
    50                  55                  60      62
```

<210> 166
<211> 92
<212> PRT
<213> Homo sapiens

<400> 166

```
Met Cys Thr His Met Gly Phe Thr Gly Lys Ser Thr Trp Lys Trp Glu
  1               5                  10                  15
Arg Thr Gln Arg Leu Pro His Leu Arg Cys Arg Ser Val Phe Thr Glu
            20                  25                  30
Gln Trp Ala Gly Ala Arg Leu Ser Ala Arg Ser Trp Glu Arg Tyr Arg
        35                  40                  45
Arg Asn Thr Gly Lys Ser Met Gly Thr Glu Gln Glu Gly Lys Gln Tyr
    50                  55                  60
Pro Gly Gly His Ser Pro Leu Ala Gln Lys Lys Ser Ile Tyr Ser Gly
 65                  70                  75                  80
Glu Gly Arg Thr Glu His Gly Gln Ser Leu Arg Ser
                85                  90      92
```

<210> 167
<211> 45
<212> PRT
<213> Homo sapiens

<400> 167

```
Met Asn Leu Glu His Val Ile Val Ser Leu Leu Thr Phe Tyr Arg Val
  1               5                  10                  15
Leu Leu Tyr Lys Glu Ile Ile Gly Leu His His Cys Phe Gln His Phe
            20                  25                  30
His Val Asn Ala Phe Leu Leu Ser Pro Leu Pro Pro Ser
        35                  40                  45
```

<210> 168
<211> 79
<212> PRT
<213> Homo sapiens

<400> 168

```
Met Ser Ala Leu Trp Leu Val Tyr Val Ser Cys Gly Cys Ser Cys Ala
  1               5                  10                  15
Leu Cys Val Leu Val His Ala Gln Thr Cys Thr Ser Phe Ser Ile Leu
            20                  25                  30
Val His Ala Gln Ala His Val Leu Ser Cys Asp His Ile His Thr Ser
        35                  40                  45
Lys Cys Thr His Ala Leu Ala His Val Tyr Thr Arg Tyr Val His Pro
    50                  55                  60
Glu Val Cys Ile His Ser Cys Ala Asp Val Cys Thr Pro Glu Gly
 65                  70                  75                  79
```

<210> 169
<211> 68
<212> PRT
<213> Homo sapiens

<400> 169
Met Lys Lys Lys Glu Arg Gly Tyr Val Leu Glu Thr Gln Ile Leu Leu
1 5 10 15
Gly Phe Gln Ile Leu Ala Val Asn Lys Leu Thr Ser Leu Val Asn Ala
20 25 30
Tyr Ala Phe Pro Asp Phe Asn Gln Arg Ile Glu Phe His Thr Leu Lys
35 40 45
Gly Leu Gly Glu Lys Lys Gln Gln Thr Lys Pro Lys Ser Leu Gly Asn
50 55 60
Gly Lys Lys Glu
65 68

<210> 170
<211> 11
<212> PRT
<213> Homo sapiens

<400> 170
Met His Ile Arg Ile Ser Met Gln Val Leu Ser
1 5 10 11

<210> 171
<211> 60
<212> PRT
<213> Homo sapiens

<400> 171
Met Pro Asn Arg Ile Glu Leu Ser Ala Leu Phe Arg Val Pro Phe His
1 5 10 15
Val Asp Trp Leu Cys Tyr Val Leu Gly Phe Cys Leu Pro Ser Thr Val
20 25 30
Lys Thr His Ser Leu Ile Val Gly Ala Val Phe Gln Ser Glu Ile Val
35 40 45
Thr Asn Val Cys Gln Ile Tyr Lys Cys Lys Thr Thr
50 55 60

<210> 172
<211> 123
<212> PRT
<213> Homo sapiens

<400> 172
Met Lys Ala Arg Ser Asp Arg Thr Gly Arg Glu Leu Gly Ser Ser Val
1 5 10 15
Ser Thr Pro Arg Ala Leu Ser Arg Arg Cys Glu Gly Trp Phe Phe Ser
20 25 30
Gln Gly Thr Ala Val Thr Ser Ala Ile Glu Ile Gly Pro Glu Arg Lys

```
         35                  40                  45
Pro Ser Arg Glu Tyr Asn Val Arg Gly Gly Ala Trp Gly Met Gly Met
     50                  55                  60
Arg Ala His Gly Pro Arg Ser Ala Gly Glu Gly Cys Lys Ser Ser Ala
 65                  70                  75                  80
Ala Pro Val Asp Arg Gly Glu Gly Ile Gly Pro Ser Ala Arg Gly Arg
                 85                  90                  95
Val Cys Val Val Lys Arg Arg Arg Lys Thr Gly Arg Ser Ser Glu Gln
            100                 105                 110
Ala Glu Ser Gly Leu Gly Gly Glu Gly Ile Arg
        115                 120         123
```

<210> 173
<211> 56
<212> PRT
<213> Homo sapiens

<400> 173

```
Met Pro Leu Leu Tyr Ile Ile Cys Leu Arg Gln Leu Val Leu Phe His
  1                 5                  10                  15
Ser Lys Cys His Ser Gln His Ser Cys Arg Ala Gly Gly Ile Gln Tyr
            20                  25                  30
Ser Met His Val Ser Leu Phe Leu Ser Ser Pro Ile Asn Tyr Asp Asn
        35                  40                  45
Gly Phe Leu Val Ser Pro Thr Phe
     50                  55  56
```

<210> 174
<211> 97
<212> PRT
<213> Homo sapiens

<400> 174

```
Met Pro Pro Lys Tyr Leu His Arg Thr Leu Ala Thr His Ser Ser Pro
  1                 5                  10                  15
Pro Val Pro Pro Pro Val Phe Ser Asn Gln Gln Met Val Pro Ser Ala
            20                  25                  30
Ala Gln Gly Trp Ser Leu Gln Pro Glu Gly Thr Leu Asp Ser Ser Leu
        35                  40                  45
Ser Leu His Ser Gln Pro Ile Lys Gln Ile Leu Leu Ile Pro Pro Pro
     50                  55                  60
Lys Tyr Ile Cys Lys Ser Cys Cys Ile Leu Cys Cys Ile Ser Leu Pro
 65                  70                  75                  80
Leu Tyr Leu Pro Glu Ile Leu Ser Ile Leu Arg Asn Arg Arg Pro Arg
                 85                  90                  95
Asn
 97
```

<210> 175
<211> 310
<212> DNA
<213> Homo sapiens

<400> 175

```
gattggtcca tttcatcaaa gttggggaat ttatgtgcat agagttatgc ataatattcc     60
```

ccgtagtgat gtcttctcct tcattcttca tattagtgtc tgaaaacaat aacaatgatt 120
tatttctcat tctgatggtt cttctcctgg tctctcttgg gcttactctg gctgtattgg 180
tagattggct acagctggag gatctaaaaa tggcctcacc catatgtttg gggccttgat 240
ggggatgact ggagtagctg agacttgtct ctccctggat tcattgggaa ccaaaatgat 300
ctgccatgtg 310

<210> 176
<211> 5291
<212> DNA
<213> Homo sapiens

<400> 176
tttcgtggat ctgataaatg cctgtagtca ttatggctta atttatccat gggttcacgt 60
cgtaatatca tctgattctt tagctgataa aaattataca gaagatcttt caaaattaca 120
gtctcttata tgtggtcctt catttgacat agcttccatt attccgttct tggagccact 180
ttcagaagac actattgccg gcctcagtgt ccatgttctg tgtcgtacac gcttgaaaga 240
gtatgaacag tgcatagaca tactgttaga gagatgcccg gaggcagtca ttccatatgc 300
taatcatgaa ctgaaagaag agaaccggac tctgtggtgg aaaaaactgt tgcctgaact 360
ttgtcagaga ataaaatgtg gtggagagaa gtatcaactc tacctgtcat cattaaaaga 420
aacattgtca attgttgctg tggaactaga actgaaggat ttcatgaatg ttctcccaga 480
agatggtact gcaacatttt tcttgccata tcttctctat tgcagtcgaa agaaaccatt 540
gacttaaagg tatcatttga aaaataccat aatggcattt gagactgaat ttctaaaaat 600
tgaatgccaa agtacaagta gaggagtttt ttattttata tatcacacac acacacacac 660
acacacacac acacacacac atatatgata caaatgcttt caggctgctt accttaccgt 720
gtagtggtaa ctattcactt cttaatttat gacctcaatc aatttaattg tctagaatgt 780
aaaaagtctt taagacataa gaattcctca aagaagccat acatttttta aggtggggat 840
tgacttttat tccaaggaac aacatcagtt cactgttgtt ggagacatga caatcatttt 900
catcccaaga acactttaag gaaacatttt acaagtatgc ttgaaagaat gtcactaact 960
ggtccagaat tttatcttct tgatttttcc agatttctct atgtttttga gaaagatgtt 1020
aatgttttgc catggtaaaa gatttcaaac cctcattttt tttgttccct tttccttgtt 1080
acttttagg aaaaactcat gctctgtttc tctgaatcaa atgaagtaga agtttacaaa 1140
gctaactttc ttcttgtcta gctattaaca tgatttgtca aatgcatgtt tttttcagcc 1200
aaagccttgt ttccattttt gttgatgtgt actcttgctc ttttagctag agtgtatgtg 1260
aaaataaaga aatatatcat tgtattcaca accatgtgtc ttcatttata acttttgtt 1320
taaaaaattt ttagttcaag tttagttcat tgatattatc ctctgaatgc agttaaggct 1380
gggcagaaat tctactcatg tgacatctgc cacaggtcta ttttgaagct tttcttctaa 1440
tggcaatgtt tgtccttacc aggatttaat ctatagaatt gtctctcaac tctgcttttc 1500
tccagttcca gataacgtcc ttaagaccat ctgttcaggg gttcacaaaa ctcaaatttg 1560
tgtcattcta ttttatttat tttatttttt atttccttcc ctcatacctt gcccattccc 1620
tttgaatatt aggtgtgatg tcaacagcat gttagaagga tcaatgggaa ggcaatgatt 1680
gaaaacattt caatgaacct taatagtgtt cctttgagga gcacccagga gaatatctgg 1740
tcatagatct ttttttaaat gcagttttat aaaaccctaa cagcggtgat atcattagac 1800
tgtatgaatc agttttatta cctagtgtac aagtgtcagt catgtatcat tatatagtct 1860
gttgatcttt ccatttgcaa aaaattaata gttttccccc acaaatgtac aaagttgtat 1920
gcttccagtc ttcttttaat gtttatagtc attccaaagt aacattctat tttacacttt 1980
cacatacatt gttatgaatc attggttttt ctctttttc cacttatcac caatttattt 2040
cattcagcca gatttggtgt cttcattttg tagaacggtg taagtggttt ccattccagc 2100
atgaatgtgg tcggtcacat ggcagtggag taaccaaatt ccaggtgttc ttggaaacat 2160
ttctagggtt tggtatgttc cagggaaaat gtcaaagaca tcagaactat aaactccccct 2220
gtgcttgtat tggaagctat ggccgtgaaa atgtacagtg tgtaagtcta tttcattgcc 2280
cattcccatc agataccagt tgacttcatc tcccacgtgc attgtgaggc cttgtaggtt 2340
tccaaacatt cttccattaa tagcatgcat tttattgctt tctatgaatt cctcatcatc 2400
tttgtttact ttctcggggt gatcagagta tgttttgatg ttgtcatcta agtaccaaga 2460
ttcattctca tcaaaaacta gaaacagaag ggcaaattcc agtttccttc tgggattgaa 2520
tactttcaag taaggtcttc gacaaacaat caggggggcca attaatccac tgtagaggtc 2580
cttaacttga tccacagttg aataataagc ccatggaata caagcagaat cctctgttcc 2640
agctccagat ctttctggga ttttccatac gtaagtgaga gtttcacctg gtaatgttgg 2700
agtaactgta gaactctctg tttgtacccc atgggcatgt attgagtagg gccttgtggc 2760
catgttttta aagataattt tgactttgtc tccaacatct gcatgaagtt gtggacctag 2820
aattcccaga tgttcttctt cagcttttct ctccactgga acacggaatg tgctatcagt 2880
atactgccga tacacaactt tcttgtactt tgagcctatg taaaactctc ccttatctaa 2940

aaatgcattt gaaacattct gctcttgtaa atgatgcagc tccttttccc actcccttg 3000
tggggaataa tcccattcca cctccactgc tgcgatatag tatgtcctct ctcccaggta 3060
gaaggtggaa tcctcagact gccgcctgca ttggttcaca gtatattttt gcttcatgcc 3120
gcctgtgtaa tgatcagttg taaggcattc aacattaaaa gtccccctctg tgtcaggcca 3180
catgtggagc gtaagacttg tttgagggaa gaggtttgct gtgtctctcc gttctcctct 3240
ccacagatat gtgtttcctg aaaagtatat tccatgtaca tcggcctcat ttccggcgct 3300
gaataagtac cacacgaccg aatctccttt gcacatagtg agacccggct gattcccata 3360
catgaatcca ttcatggagt gcattttatt agattcctga aagtcttcat cttccttatc 3420
cacctgatca ggtgcagttg taaacattct aatattatct tccaggagta aactctcatt 3480
ctcatcaaat actgtaggaa acaaatagaa ttccttgtct acatctttct gtctcccatt 3540
tgcatgtaaa cttcctttct tgcatatttt cattggccca ataagcccag tgaaatatat 3600
cctttagtgg gatccacagc agaataatac atcttagcta gacacacagg atctgcatta 3660
gtgggtccta cttctttggg gacagtccat tcataggtga atgtttctgt gggtgccaca 3720
tgggaggctg aaggaggcac acttctgctc tgggggttgt aatttgggga atagtatgtg 3780
ccctcgttgt tcttattgaa tctcacccca atcggctcaa tactgagggg atatgctcct 3840
ttgttatgga aggttactct gatggtgtct cccacctctg cccaaatgac aggacccagg 3900
atgccaagat gctcttcttc agggcctctc tcctttcgat ttgtgaagga ggcatctgtg 3960
tactcacgat aaaccagctt tttataagag cctccaattc ttgtggtacc ttgttcaaaa 4020
aaacaccgct gagtcacttc caggtgctgt taagttttct ttagtgaaga tgtctatacc 4080
agagggagca tagttccaga tgatttcctc agcggcaatg tagtagtgtc taacatgctt 4140
cccacggata ttatcctttg atgaagactt gttacactcc tggacctgga aaaaggcttg 4200
caaaccggct ttcagatggt ttagattctg acagctgagc atccattctc cagggttctg 4260
ggccaccata taagcatcaa acagggtagc aggaaagagg ttgattgtgt caatacggta 4320
gttcttgtta gtcagtgctt gcccgtgaaa gaaagctgcg tgcacatcaa cttcattacc 4380
cataccaaaa aggtaccatt ttactctgtc ttcagcacac atggagagtc ctgggagact 4440
tccaaaagtg tatccattca cagaatacat tctgttactc tcctggaagt cttcgttgtc 4500
tttgtcaact ttctctggtt ctgagcagta ggttttaatg ttgtcttcta ggtaccagct 4560
gaaattttca tccaccacag aaaacatcac cacaaattct cggtcaatat gtttttcttt 4620
ttctttatct agagaatctt ttttacagat tattaaaggt ccgatgagtc ctgaggcaat 4680
atcttttgga gcatcaatgt gggaatggta aatcctagtc acacaattgc catctccttc 4740
cccaggactt tgttcttcag tggcaagcaa catgtatgta tactgctctc ctggatatac 4800
tttgtcatct gctctttgaa aatctgtggt gttatcaggg tagatggccc cctcatgttc 4860
cttatagtaa gttattccat gtgaatgaaa ggtgtggggc ctaggaggca aggtttttta 4920
agtgtacata aactttatct ccagtttcag ctttgataat agggcctaaa aacccaagcc 4980
agaccggttt ttctatagtt gtcctaaagg tttcatctgt gtactgaaga taaagggcct 5040
tcttatatag tctcccaatt ctatctgggc cattttgaag atagatattg gaatgttccg 5100
tgtcaacaga aataagtttc ttttccccat ggtcagaggc ataatcccaa gtcgtttcaa 5160
taattccaat gtaataatgc ttttctttcg cccaggctgg ggtactacat aaaaacagaa 5220
aaataccaag tatcaaaatc ttcatttttt tccccttctt ggcggacgcg tgggcggacg 5280
cgtgggtcga c 5291

<210> 177
<211> 2186
<212> DNA
<213> Homo sapiens

<400> 177
ggggcctttt cttgctgaat gactcaccag tgatctttga atgcagacaa gcttctagcc 60
ctccctcttc attagtgagc atttctatac aggggcctat atagtgcttg gagtctagga 120
gctatacaat tgggggggaaa agaagcccta cagattctgg cctctactta ggaagtgctg 180
aaggactctg agtgacttta ttcattaaaa caggttgcta ttaattaact gtcataataa 240
aataaagaac tattatttgc atttcttact atgttctctt tagtaatatg tataaaagat 300
gcttgtaaca gaaatgaatt tattgggcca cttccccatc ccccccaatc cctgcctttt 360
tttccccttt gcattttggc agcaggagcc tactaattat tgggtctaga tctctgagtc 420
caaatagaaa taaaacataa atataaaatt aattaagtaa atgcaaaatc ttttttactc 480
gccagtttgt gcaactgaaa taagcaaaca tcaaattaac tttttttttt tttagtgatt 540
tccaagaacg atttaagaca gtaaaaccca aacttttttc aaaatatcac ccatttggtg 600
tttagcttga agtgtaagac tatgtaacaa ctatagctac tttgtaaaaa atgtactctc 660
acctaaagcc taacatgatt catttttttt tttccaggaa acagaaaatg tctttggtct 720
ggtattaata tcaaaagaat ttaactttaa aactgatccc aataagcatg gcccttacac 780
agttcatgac caatatttag ttatatgaaa atggttcatt ctcttctaga atctaatcca 840

gtgcagtcat ctggtgaaag actgggaagt cagaagactt ttcaaaagct tctcactgaa 900
atgggtaatt tgggtaacac gggataaaca gaattttgca taaatagcag accttagtca 960
ctgaataccc cattaaacag tactgctgga ccatgatact tatatctggg agatgactta 1020
gttttaaatt ttttatgcag tttcttttct ttgaaggaaa aacttccaga cttcagaatt 1080
ttggaagctg aacgatgagc aggtttttca tattttgctg tctcagacat ttttcctatt 1140
tctcagactt tgcagttctt ttcttaggtg ctttagagca cctaaaatat cagcttgctg 1200
tgggtcattc ggtgctttct gagagtacag acggaaatgt ctgatggttt ccgataacag 1260
cacttcagga tctgcatctc ttgcctgaag acgcagcatc cgctcacagg cgatggcata 1320
gttcttgtgc caattcactg ggtgctcctt ttgtaaatag acaatctcct tataaaggtc 1380
atatgcttga gcgtagagtt gagctttcac ttttgaaggc atgtgagcag tttctgccag 1440
gttaaagatg aagaacggtg ttttcatcct cgcttgccac atttcttcat tggccacaga 1500
ttcccaagaa gatggatcaa accttccata ttcttcggtc cagttataga tatttttgt 1560
aagtttaatc cattcctcag ggttgaatac aatctccaaa ggaactaatt tgtcacaaga 1620
cccccatggc caaagtgaat agttcttttt ccaggttggg tcgccttcat gaattcctat 1680
gcaaacaaat gtttcttttt gtttatttac ttcaagaaaa tgataaagat taaatgtgac 1740
cattccacta ggtaatattc cttccacagg attccaccgg ttcccaggaa agttgacacc 1800
tggcaagtgc tttgccatct tgggtaaata ccactcgtaa gtcatcattt cctgatccac 1860
taatgaaatg tcaggcctca acccctcaca gtaatgcatg taacggagag aatttcctgg 1920
caaatctcct ctgagtaaga taattgcatc atgaggcata gaggtgagaa ggttctttgc 1980
gaacttatca atcacatagt tggtcctttg gtcacaaaca ctgtaattag aatatatttg 2040
gtaaactaca aaaagagttg cagaaagcca ttccagacac tgaagcccat tgctattcag 2100
cactcggtta gtctcagaca caactgcagc caaaccaatg ccagcgagga aggccactac 2160
tgcattgctc tgcatccaga atcgtt 2186

<210> 178
<211> 1191
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(1191)
<223> n = a,t,c or g

<400> 178
tttcgtggaa cctgtgatcg agttacataa agacacacgt aatcataact gagcaaagct 60
cttctctaat atgtaggtga aaaaattaac tgactctgtt aatctggaaa atgattgaat 120
gtgtatttgt taaattttaa tggaaatatc caaaaaacca actcagacaa gtggtgaaca 180
gagtagccgt ggagtattta tttcttctgt atatgtatat tgcaattgag tgcagaattt 240
gggatacatg gattcctttc cttcttcctt tgtgttctct ccccagtatg ggaagactct 300
ctatgataac taccagcggg ctgtggcaaa ggctggactt gcctctcggt ggacaaactt 360
ggggacagtc aatgcagctg caccagctca gccatctaca ggacctaaaa caggcaccac 420
ccagcccaat ggacagattc cccaggctac acattctgtc agtgctgttc tccaagaggt 480
ccagagacat gctgaaacat cgaaggtaaa accagcaagc agctgacccc tataagccat 540
gttctaaacc attcacatta ttccagtaac tgctcttta aaaagagaaa ataagctgag 600
tgcagtggct cacacctgta atcccaaaac tttgggaggc tgaggcagga agatcacttg 660
agcttcagtt tgagaccaga ctgggcaata tggcgagagc ttgtctctac tanaaaaata 720
aaattaccca ggcatggtgg catgcacctg gtttctccgc tactggggag gctgaggaat 780
ttgacttgtt ttacccacga aggttaagcg gcagtgatct gtgattgcgc cctggatttc 840
accctggccg aagagtagaa ccctcttaaa taattaccta ttcgtcataa ttgaaatgcc 900
gtcccacact gaaagtactc acattattta gtggatgtgg aattggaacc taaagttact 960
ttgtggatca caagtagaac ccgtttcttc actggttaag aaaatgaagc attccctata 1020
acttgggtga ataacttttt actttacaca gaggtttctg tttaatcaag aatttaccct 1080
cttgtaaacc ctacaacatt aaccagggtg tttcggcgct taaaaacaac ttagcccggt 1140
gccangcaag ttttataggc actaataaac taggcgtgtc aagccggggc c 1191

<210> 179
<211> 322
<212> DNA

<213> Homo sapiens

<400> 179
tgctaaacgt cacggtctta aaatccatgc acagaggcgt cctagacaca ctgctgaaca 60
tcacaatctt agaatccatg cacagaggcg tcctagacac actgctaaag atcgctgtct 120
taaaatccat gcacagaggc gtcctagtca cactgctaaa gatcacagtc ttaaaatcca 180
tgcacagagg catcctagtc acactgctga agatcaccat cttaaaatcc atgcacagag 240
gcgtcctgga cacactgcta aagatcacca tcttaaaatc catgcacaga ggcgtcctgg 300
tcacactgct gaagatcaca at 322

<210> 180
<211> 1617
<212> DNA
<213> Homo sapiens

<400> 180
ccatggttat gttggctgcc ctggctcacc atttgtttta ctgggatgtt tggtttatat 60
ataatgtgtg tttagctaag gtaaaaggct acaggtctct ttccacatcc caaactttct 120
atgatgctta catttcttat gacaccaaag atgcctctgt tactgactgg gtgataaatg 180
agctgcgcta ccaccttgaa gagagccgag acaaaaacgt tctcctttgt ctagaggaga 240
gggattggga cccgggattg gccatcatcg acaacctcat gcagagcatc aaccaaagca 300
agaaaacagt atttgtttta accaaaaaat atgcaaaaag ctggaacttt aaaacagctt 360
tttacttggc tttgcagagg ctaatggatg agaacatgga tgtgattata tttatcctgc 420
tggagccagt gttacagcat tctcagtatt tgaggctacg gcagcggatc tgtaagagct 480
ccatcctcca gtggcctgac aacccgaagg cagaaggctt gttttggcaa actctgagaa 540
atgtggtctt gactgaaaat gattcacggt ataacaatat gtatgtcgat tccattaagc 600
aatactaact gacgttaagt catgatttcg cgccataata aagatgcaaa ggaatgacat 660
ttctgtatta gttatctatt gctatgtaac aaattatccc aaaacttagt ggtttaaaac 720
aacacatttg ctggcccaca gtttttgagg gtcaggagtc caggcccagc ataactgggt 780
cctctgctca gggtgtctca gaggctgcaa tgtaggtgtt caccagagac ataggcatca 840
ctggggtcac actcatgtgg ttgttttctg gattcaattc ctcctgggct attggccaaa 900
ggctatactc atgtaagcca tgcgagcctc tcccacaagg cagcttgctt catcagagct 960
agcaaaaaag agaggttgct agcaagatga agtcacaatc ttttgtaatc gaatcaaaaa 1020
agtgatatct catcactttg gccatattct atttgttaga agtaaaccac aggtcccacc 1080
agctccatgg gagtgaccac ctcagtccag ggaaaacagc tgaagaccaa gatggtgagc 1140
tctgattgct tcagttggtc atcaactatt ttcccttgac tgctgtcctg ggatggcctg 1200
ctatcttgat gatagattgt gaatatcagg aggcagggat cactgtggac catcttagca 1260
gttgacctaa cacatcttct tttcaatatc taagaacttt tgccactgtg actaatggtc 1320
ctaatattaa gctgttgttt atatttatca tatatctatg gctacatggt tatattatgc 1380
tgtggttgcg ttcggtttta tttacagttg cttttacaaa tatttgctgt aacatttgac 1440
ttctaaggtt tagatgccat ttaagaactg agatggatag cttttaaagc atcttttact 1500
tcttaccatt ttttaaaagt atgcagctaa attcgaagct tttggtctat attgttaatt 1560
gccattgctg taaatcttaa aatgaatgaa taaaaatgtt tcattttaaa aaaaaaa 1617

<210> 181
<211> 3592
<212> DNA
<213> Homo sapiens

<400> 181
ctcaggctca taccacacgg ggggtgctag tgggctgctg tgtgtttaca tgagtaatgg 60
agctgccggg ggaggggagt tgtaagcaga gcgctgagcc tcgcagctcg cattcggagg 120
gaagctgaca tccacaccaa gtcgagactt ccagggatgt ggccggggag cagtcacatg 180
ctgtagcttt catgagcaca ggcatcagtc aggcagatgt ttgtcgactg gaatggcgcc 240
aaatcttaaa ggcagaccac gcaaaaagaa accatgccca caaagaagag attcattcag 300
tggtgttaag gattccaaca acaattccga tggcaaagcc gttgccaagg tgaaatgtga 360
ggccaggtca gccttgacca agccgaagaa taaccataac tgtaaaaaag tctcaaatga 420

agaaaaacca aaggttgcca ttggtgaaga gtgcagggca gatgaacaag ccttcttggt 480
ggcactttat aaatacatga aagaaaggaa aacgccgata gaacgaatac cctatttagg 540
ttttaaacag attaaccttt ggactatgtt tcaagctgct caaaaactgg gaggatatga 600
aacaataaca gcccgccgtc agtggaaaca tatttatgat gaattaggcg gtaatcctgg 660
gagcaccagc gctgccactt gtacccgcag acattatgaa agattaatcc taccatatga 720
aagatttatt aaaggagaag aagataagcc cctgcctcca atcaaacctc ggaaacagga 780
gaacagttca caggaaaatg agaacaaaac aaaagtatct ggaaccaaac gcatcaaaca 840
tgaaatacct aaaagcaaga aagaaaaaga aaatgcccca aagccccagg atgcagcaga 900
ggtttcatca gagcaagaaa aagaacaaga gactttaata agccagaaaa gcatccctga 960
gcctctccca gcagcagaca tgaagaaaaa aatagaaggg tatcaggaat tttcagcgaa 1020
gcccctggca tccagagtag acccagagaa ggacaacgaa acagaccaag gttccaacag 1080
tgagaaggtg gcagaggagg cgggagagaa ggggcccaca cctccactcc caagtgctcc 1140
tctggcccca gaaaaagatt cagccttggt ccctggggcc agcaaacagc cactcacctc 1200
tcctagtgcc ctggtggact caaaacaaga atccaaactg tgctgtttta cagagagccc 1260
tgaaagtgaa ccccaagaag catccttccc caggcttccc caccacacag gccaccgctg 1320
gcaaaccaga atgagacgga ggatgacaaa ctgcccgcca tggcagatta cattgccaac 1380
tgcaccgtga aggtggacca gctgggcagt gacgacatcc acaatgcgct caagcagacc 1440
ccaaaggtcc ttgtggtcca gtcgtttgac atgttcaaag acaaagacct gactgggccc 1500
atgaacgaga accatggact taattacacg cccctgctct actctagggg caacccaggc 1560
atcatgtccc cactggccaa gaaaaagctt ttgtcccaag tgagtggggc cagcctctcc 1620
agcagctacc cttatggctc cccacccccc ttgatcagca aaaagaaact gattgctagg 1680
gatgacttgt gttccagttt gtcccagacc caccatggcc aaagcactga ccatatggcg 1740
gtcagccggc catcagtgat tcagcacgtc cagagtttca gaagcaagcc ctcggaagag 1800
agaaagacca tcaatgacat ctttaagcat gagaaactga gtcgatcaga tccccaccgc 1860
tgcagcttct ccaagcatca ccttaacccc cttgctgact cctacgtcct gaagcaagaa 1920
attcaggagg gcaaggataa actcttagag aaaagggccc tcccccattc ccacatgcct 1980
agcttcctgg ctgacttcta ctcgtccccc catctccata gcctctacag acacaccgag 2040
caccatcttc ataatgaaca gacatccaaa tacccttcca gggacatgta cagggaatcg 2100
gaaaacagtt cttttccttc ccacagacac caagaaaagc tccatgtaaa ttatctcacg 2160
tccctgcacc tgcaagacaa aaagtcggcg gcagcagaag cccctacgga tgatcagcct 2220
acagatctga gccttcccaa gaacccgcac aaacctaccg gcaagggcct gggcctggct 2280
cattccacca cagggcccca ggagagcaaa ggcatctccc agttccaggt cttaggcagc 2340
cagagtcgag actgtcaccc caaagcctgt cgggtatcac ccatgaccat gtcaggccct 2400
aaaaaatacc ctgaatcgct ttcaagatca ggaaaacctc accatgtgag actggagaat 2460
ttcaggaaga tggaaggcat ggtccaccca atcctgcacc ggaaaatgag cccgcagaac 2520
attggggcgg cgcggccgat caagcgcagc ctggaggatt tggaccttgt gattgcaggg 2580
aaaaaggccc gggcagtgtc tcccttagac ccatccaagg aggtctctgg gaaggagaag 2640
gcctctgagc aggagagtga aggcagcaaa gcagcgcacg gtgggcattc cgggggcgga 2700
tcagaaggcc acaagcttcc cctctcctcc cctatcttcc caggtctgta ttccgggagc 2760
ctgtgtaact cgggcctcaa ctccaggctc ccggctgggt attctcattc tctgcagtac 2820
ttgaaaaacc agactgtgct ttctccactc atgcagcccc tggctttcca ctcgcttgtg 2880
atgcaaagag gaatttttac atcaccgaca aattctcagc agctgtacag acacttggct 2940
gcggctacac ctgtaggaag ttcatatggg gacctttttgc ataacagcat ttacccttta 3000
gctgctataa atcctcaagc tgcctttcca tcttcccagc tgtcatccgt gcaccccagt 3060
acaaaactgt aggctcagct ctgcccagca gtccaaagcg gcatggccaa cagagcttca 3120
ctccttaccc aggagtgctg gcttatagag ttagaagtca gtatttcttc taatctgagg 3180
ctatgatcag tcccagctgt aggggcccag aggggaggtg aacatgcctg atttttgtgg 3240
gacaactgta gcccacaaac tgactggctg gtgagtcttg actcccttcc aacacagatg 3300
cccaggcacc tccagatcat tcacttcgca cgtgggcctt gtgaagggat ttgtgaatat 3360
ccaggaagaa cttagaggac cccatctgag ttcggatggt caggaaacaa tctgggcaaa 3420
aaagaggcag gcatttcaaa ggaaggggca aggaagactg gcaaacagat ggcaagggat 3480
gcccctcttt ttcataaaac tctccaaggt tcaatcaatg caatgtatag tgaaacttca 3540
atagatcttt cattttgaca ctattaaaca atccagagaa gtaaacactg tt 3592

<210> 182
<211> 1256
<212> DNA
<213> Homo sapiens

<400> 182
ggcacgaggc aggaagtgac tgcgggagtg gagccggcga gagagtggca gcgggggctg 60

atggaagtgc agtgggggct ggagagggca ccctactgta tccagcatgc tccaaggcca 120
cagctctgtg ttccaggcct tgctggggac cttcttcacc tgggggatga cagcagctgg 180
ggcagctctc gtgttcgtat tctctagtgg acagaggcgg atcttagatg gaagtcttgg 240
ctttgctgca ggggtcatgt tggcagcttc ctattggtct cttctggccc cagcagttga 300
gatggccacg tcctctgggg gcttcggtgc ctttgccttc ttccctgtgg ctgttggctt 360
caccettgga gcggcttttg tctacttggc tgacctcctg atgcctcact tgggtgcagc 420
agaagacccc cagacggccc tggcactgaa cttcggctct acgttgatga agaagaagtc 480
tgatcctgag ggtcccgcgc tgctcttccc tgagagtgaa ctttccatcc ggataggtag 540
agctgggctt ctttcagaca agagtgagaa tggtgaggca tatcagagaa agaaggcggc 600
agccactggc cttccagagg gtcctgctgt ccctgtgcct tctcgaggga atctggcaca 660
gcccggcggc agcagctgga ggaggatcgc actgctcatc ttggccatca ctatacacaa 720
cgttccagag ggtctcgctg ttggagttgg atttggggct atagaaaaga cggcatctgc 780
taccttttgag agtgccagga atttggccat tggaatcggg atccagaatt tccccgaggg 840
cctggctgtc agccttccct tgcgaggggc aggcttctcc acctggagag ctttctggta 900
tgggcagctg agcggcatgg tggagcccct ggccggggtc tttggtgcct ttgccgtggt 960
gctggctgag cccatcctgc cctacgctct ggcctttgct gccggtgcca tggtctacgt 1020
ggtcatggac gacatcatcc ccgaagccca gatcagtggt aatgggaaac tggcatcctg 1080
ggcctccatc ctgggatttg tagtgatgat gtcactggac gttggcctgg gctagggctg 1140
agacgcttcg gaccccggga aaggccatac gaagaaacag cagtggttgg cttctatggg 1200
acaacaagct tctttcttca cattaaaact tttttccttc ctctcttctt catctc 1256

<210> 183
<211> 358
<212> DNA
<213> Homo sapiens

<400> 183
tgagaatgca actaatctta gttaacaaaa gaatcataga gttaaaggag atcttgatta 60
tttctttcta ttcatccttt tgactttgtg taggatgata ccaacaacat tccagaaaaa 120
cttcattctg agcttttgca ttcaccccaa atttctcaga tgaaatgtaa gcttatccct 180
gtttgtccct ttctcagatt gaatacacag ccgctgctca ttatctctta tgggatattc 240
ttgcatatat ttcgagactt tagctatatt cacagggtca gagagaggca tagtgtcttc 300
ctatctgtag gacagcagtg gtgtcctgaa ttgacaagga gcattttctt attgaatt 358

<210> 184
<211> 3320
<212> DNA
<213> Homo sapiens

<400> 184
ggcacgagcg gcacgaggcg gcgggagggg aaacgtgcgc ggccgaaggg gaggcggagc 60
ccgcgccggc tgcgcagagg agccgctctc gccgccgcca cctcggctgg gagcccacga 120
ggctgccgca tcctgccctc ggaacaatgg gactcggcgc gcgaggtgct tgggccgcgc 180
tgctcctggg gacgctgcag gtgctagcgc tgctgggggc cgcccatgaa agcgcagcca 240
tggcggcatc tgcaaacata gagaattctg ggcttccaca caactccagt gctaactcaa 300
cagagactct ccaacatgtg ccttctgacc atacaaatga aacttccaac agtactgtga 360
aaccaccaac ttcagttgcc tcagactcca gtaatacaac ggtcaccacc atgaaaccta 420
cagcggcatc taatacaaca acaccaggga tggtctcaac aaatatgact tctaccacct 480
taaagtctac acccaaaaca acaagtgttt cacagaacac atctcagata tcaacatcca 540
caatgaccgt aacccacaat agttcagtga catctgctgc ttcatcagta acaatcacaa 600
caactatgca ttctgaagca aagaaaggat caaaatttga tactgggagc tttgttggtg 660
gtattgtatt aacgctggga gttttatcta ttctttacat tggatgcaaa atgtattact 720
caagaagagg cattcggtat cgaaccatag atgaacatga tgccatcatt taaggaaatc 780
catggaccaa ggatggaata cagattgatg ctgccctatc aattaatttt ggtttattaa 840
tagtttaaaa caatattctc tttttgaaaa tagtataaac aggccatgca tataatgtac 900
agtgtattac gtaaatatgt aaagattctt caaggtaaca agggtttggg ttttgaaata 960
aacatctgga tcttatagac cgttcataca atggttttag caagttcata gtaagacaaa 1020
caagtcctat ctttttttt ttggctgggg tgggggcatt ggtcacatat gaccagtaat 1080

tgaagcacgt catcactgaa agacagaatg ccatctgggc atacaaataa gaagtttgtc 1140
acagcactca ggattttggg tatcttttgt agctcacata aagaacttca gtgcttttca 1200
gagctggata tatcttaatt actaatgcca cacagaaatt atacaatcaa actagatctg 1260
aagcataatt taagaaaaac atcaacattt tttgtgcttt aaactgtagt agttggtcta 1320
gaaacaaaat actccaaaaa aaaaaaaatt tttcaaataa aacccaaaat taatagcttt 1380
gcttaggccc tgttagggat ccattgggag cattaaggga gcacatattt ttattaactt 1440
cttttgagct tttcaatgtt gatgtaattt ttgttctctg tgtaatttag ggtaaactgc 1500
agtgtttaac ataataatgt tttaaagact tagttgtcag tattaaataa tcctggcatt 1560
atagggaaaa aacctcctag aagttagatt atttgctact gtgagaatat tgtcaccact 1620
ggaagttact ttagttcatt taattttaat tttatatttt gtgaatattt taagaactgt 1680
agagctgctt tcaatatcta gaaattttta attgagtgta aacacaccta actttaagaa 1740
aaagaaccgc ttgtatgatt ttcaaaagaa catttagaat tctatagagt caaaactata 1800
gcgtaatgct gtgtttatta agccagggat tgtgggactt cccccaggca actaaacctg 1860
caggatgaaa atgctatatt ttctttcatg cactgtcgat attactcaga tttggggaaa 1920
tgacattttt atactaaaac aaacaccaaa atattttaga ataaattctt agaaagtttt 1980
gagaggaatt tttagagagg acatttcctc cttcctgatt tggatattcc ctcaaatccc 2040
tcctcttact ccatgctgaa ggagaagtac tctcagatgc attatgttaa tggagagaaa 2100
aagcacagta ttgtagagac accaatatta gctaatgtat tttggagtgt tttccatttt 2160
acagtttata ttccagcact caaaactcag ggtcaagttt taacaaaaga ggtatgtagt 2220
cacagtaaat actaagatgg catttctatc tcagagggcc aaagtgaatc acaccagttt 2280
ctgaaggtcc taaaaatagc tcagatgtcc taatgaacat gcacctacat ttaataggag 2340
tacaataaaa ctgttgtcag cttttgtttt acagagaacg ctagatatta agaattttga 2400
aatggatcat ttctacttgc tgtgcatttt aaccaataat ctgatgaata tagaaaaaaa 2460
tgatccaaaa tatggatatg attggatgta tgtaacacat acatggagta tggaggaaat 2520
tttctgaaaa atacatttag attagtttag tttgaaggag aggtgggctg atggctgagt 2580
tgtatgttac taacttggcc ctgactggtt gtgcaaccat tgcttcattt ctttgcaaaa 2640
tgtagttaag atatacttta ttctaatgaa ggccttttaa atttgtccac tgcattcttg 2700
gtatttcact acttcaagtc agtcagaact tcgtagaccg acctgaagtt tctttttgaa 2760
tacttgtttc tttagcactt tgaagataga aaaaccactt tttaagtact aagtcatcat 2820
ttgccttgaa agtttcctct gcattgggtt tgaagtagtt tagttatgtc tttttctctg 2880
tatgtaagta gtataatttg ttactttcaa atacccgtac tttgaatgta ggtttttttg 2940
ttgttgttat ctataaaaat tgagggaaat ggttatgcaa aaaaatattt tgctttggac 3000
catatttctt aagcataaaa aaaatgctca gttttgcttg cattccttga gaatgtattt 3060
atctgaagat caaaacaaac aatccagatg tataagtact aggcagaagc caattttaaa 3120
atttccttga ataatccatg aaaggaataa ttcaaataca gataaacaga gttggcagta 3180
tattatagtg ataatttttgt attttcacaa aaaaaaagtt aaactcttct tttcttttta 3240
ttataatgac cagcttttgg tatttcattg ttaccaagtt ctatttttag aataaaattg 3300
ttctccttct aaaaaaaaaa 3320

<210> 185
<211> 435
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(435)
<223> n = a,t,c or g

<400> 185
cggacgcgtg gggcagaagg tgcaagccag agctcaggca gaacttccag attgcatctg 60
ggatctgcat ttgccactgg ttgcagatca ggcggacgag gagccgggaa ggcagagcca 120
tgtggctgcc ccctgctctg ctccttctca gcctctcagg ctgtttctcc atccaaggcc 180
cagagtctgt gagagcccca gagcaggggt ccctgacggt tcaatgccac tataagcaag 240
gatgggagac ctacattaag tggtggtgcc gaggggtgcg ctgggataca tgcaagatcc 300
tcattgaaac cagagggtcg gagcaaggag agaagagtga ccgtgtgtcc atcaaggaca 360
atcagaaaga ccgcacgttc actgtgacca tggaggggct caggcgagat gacgcanatg 420
tttactggtg tgggg 435

<210> 186
<211> 827
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(827)
<223> n = a,t,c or g

<400> 186
aaaaattgtg agaatgaaat tccaccttag tttctttttct ctgaaaagag ccatatttta 60
tatatgtgcc aaagctgata aaatcagtgg aggctatttg tacaaatgta ggacagtgtc 120
gtatagtgga aagaatgtta gatctggagt taaaatatct ggatttctgt ctgcttgcat 180
cattagctat ctgtgattgt gagcaagttt cctagtgttt taagtaattc acccaagatt 240
atgcaagtag tgagttgcag aaccaggatt cagatgcaaa tctttctaac cctagggctc 300
tatctctagt gtctatactc ctaacctaga ctgtggtgtt tcccctaact tctgaatctt 360
tctgatacag attttgtccc catgaaaggg aacattctca cacccacttg gtttcatttg 420
ttgctagtaa agataaaatg ggatcatacg taggtaaagc ccactgtgca gtctacatta 480
tcatgatact tattgtataa tatctgatta ttcactattc aaagctctac ctctgagaac 540
taggtataaa tgaagttagc tcttgagcgt ctgaaatcaa ctaacaaaca cagcccactc 600
attggagaca ttcattcaac actatttact tatttttttaa aagataaatg cattgtgtat 660
atttgaggcg tataacatat ngtatggggt ttatatataa ataagaaaaa tggttagtga 720
agnaaattaa catggctggc atctcacata ggtacttttc ttgggggaag gcaaaaggca 780
gctaaaaatc tgcctatttt aacaaaaaaa cctagtacat ttacatc 827

<210> 187
<211> 689
<212> DNA
<213> Homo sapiens

<400> 187
tttcgtgggg aaccgagcgc ccggcgcgga gagcgcggga gcggagcagt agcccgatcc 60
ggggccgcgg gccgagctgc cgatctttct cctggcctgc cctccatgtc ggggagctat 120
cgtcgtcttc aagctacaga tgcacatgtt gaatggagct cttctggcat tgctgtttcc 180
tgtggtaaac actcggctgc tccccctttga attggagatt tactacattc agcatgttat 240
gctctacgtg gtacccatct acctgctttg gaaaggaggt gcttacactc cagagcccct 300
cagcagtttc cggtgggctc ttctctcaac tggcctcatg ttcttttatc acttcagcgt 360
cttgcagatc ctcggcctgg tcaccgaagt gaatttgaac aacatgctgt gtccggccat 420
ctcagaccca ttctacggcc cctggtatcg catctgggcc tcgggacacc agactctcat 480
gaccatgacc cacgggaagc tggtcatcct gttctcatac atggctgggc ccttgtgtaa 540
atatctgctg gatttgctcc ggcttccagc caagaaaata gactgaaggt gcttatttt 600
ttttttttc ctccctgagg aagcaagtcg tgacttgact tggagaacac ccagttcttg 660
ataaaatcat gggagagggc aaaaaaaaa 689

<210> 188
<211> 1127
<212> DNA
<213> Homo sapiens

<400> 188
gctccccgcc gaggctacgg tctttacctg tcgcgtagag aggacgcgcc gcgcgcgagt 60
attacgtagg agggttacct gcaggcgccg cgctgccgca ggctttgcaa gagtgtaaaa 120
ggggcccggg gctaggggag atacgaattt gattgcaggg ggtcgtagcg agcagtagag 180
gcgtcgataa tggggaggta accgggggggg gggggtcccg aaaggaaagt ccgttttccc 240
tgagcctgga gcaactaatc ggatttcgcg tgaggggggc gtgggaccct ttgctcccgg 300
ccgaggctgc ctgtcctgcc atcagagctt ggtctcgctg aggatgatat tggccgtgac 360

gaaaatgaag caggagaaag cccagagcag cacgaagccc acccagaagg tgtggcggaa 420
aggggacctg tgcttgttgc tgaccccaca gccaaaggcc agctgggtgt ccctgcggca 480
gtggtacacc aggaaggcgg ggatgacgta ctggatgccg gtgcccgcgt aggcccctgt 540
gatgcccacc agggactcca ggtcgtgggt gcagaaggcc accagcacag gcggcaccag 600
ggtgatggtg ggaaacacga cgcggtccac cacccacggg tacgtgccgc cctcgcggtg 660
gaagagtgtc ttccagttgt tgcgcagggt cacggcaatg atggggaagt tggtgctgat 720
ggtgaagacg gggaagaggc ccaggaaaag gcgcgcagcg gccaggccca cgacgtcaca 780
gcgcgcgaag ttgagggtgt acatgtccat gaggctgtcg ccgcggaagc agaagatggc 840
ggtgaaggag aggaggccgt agaaggccag gatcagcacg tagtccagga acaccagcct 900
tgtgaggtgg cgcttggagg agacgggggt aatgagggat ggcagagagt gctggcacat 960
gaaggagtag acgcacaccc caaacaggtt ccggaccccc gagaagtcag ccaggggcgg 1020
gtgccctcc ccttgtccgt gcccgatgcg gatcagggcc agcacaatca tgacggcgaa 1080
agcgatccat ctcatcagag aggtcaggat ctgcaggtac tacgaaa 1127

<210> 189
<211> 372
<212> DNA
<213> Homo sapiens

<400> 189
cgacgctagg aactgggcat tggagatcag aggcatggca ctggctttgg ctgcctatgt 60
atgtggctgg gtggtggaca gggaaacatg gcctgtgccg atgccctgta ataaaggagg 120
tagggcttgt aacctggaga tgggcatgga gtggctgaat ctccactgcg aggtcagtaa 180
atggcagcag cctcccagtg gggctctttg ctgctctctt gctcccctgc aatccatttt 240
cttcccagca gccaaagtaa tgtttaaaaa tggcagttgg actgtgctac tcccatgctc 300
agaattcccc attggcttcc cgtcacactt ggaataaatt cccaactact taccttgctt 360
cctccccatc at 372

<210> 190
<211> 2653
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(2653)
<223> n = a,t,c or g

<400> 190
gttccccttc ctttcctctg tccaattccc tgagaaagtg gctgacagag ccccccccaga 60
agtagctgga gggacctcag gcccccgaat gcctggagac agaggggtcc cccagtcctc 120
agcctcactg gcaggcccgg aaagtgctga attcgggggt ctcaggcgag gaggacgtgg 180
aagtgggcct cttccgggtc ctggagttgc tgggggtggc agtggtgcct gagtgatcat 240
ccaaggtggc cgggtcttct tctggcatcg ttttcgcgga gacggtaaac agggagttat 300
cgaaggacat gtctgtgctg tcgaagttaa tctgggccca gcaggattct cgatctcttt 360
tcttctgtaa atttctgagt ttaacttttc tgaccaccag ggcaacagcc accagcccag 420
caaggaggaa gagggcacca gccaggctca gggagagcac caagatccaa gtgggcactg 480
ggaaggagac gttcacgggc tcgctgaggt ctgacagctg ggactggttg agctcaccta 540
acactccata ctggcagtgg aagggtcccc ctggagcctt gctgctgccg ccgctcaggt 600
taaatgtcac cccgcgctgg tccgtggggg cctgcaggag ctggaccacc tgccccccctc 660
gatacagtgt gaaattcgcc cccgggaagt tcccaggggc catgcatgcg atgtggatgg 720
ggtcctcttg gctgcttggg tacgggggca ccagccggat ggatggtgct gggatcgcca 780
aggagccagc tgcaaagagc aagatggtcc agggcatctc tctctgcgtt tggctgggaa 840
atccggctga ggccagaggg gcccaaggct ccccgccaac tcctcctggt ggcccttgag 900
tctgagaccc ttggggtgca tggtggggcc acgccccctcc actgtctata agccaggaaa 960
gctgagcact caatatgaac caagcctcat ctcaattctc acagccgcac tacagggtgg 1020
tcgttccatc actcccgact tacgaggaaa ctgatacgct gatgcctcta aggcgcaaaa 1080
ggcgtcgaac ctggaccacg ctcacagagc aagcaaaaca cgacgccaca gaccagaggg 1140

tgacccggat gatggcggcc ggcgcggccc tagccctggc cttgtggcta ctaatgccac 1200
cagtggaggt gggaggggcg gggcccccgc caatccagga cggtgagttc acgttcctgt 1260
tgccggcggg gaggaagcag tgtttctacc agtccgcgcc ggccaacgca agcctcgaga 1320
ccgaatacca ggtgatcgga ggtgctggac tggacgtgga cttcacgctg gagagccctc 1380
agggcgtgct gttggtcagc gagtcccgca aggctgatgg ggtacacacg gtggagccaa 1440
cggaggccgg ggactacaag ctgtgctttg acaactcctt cagcaccatc tccgagaagc 1500
tggtgttctt tgaactgatc tttgacagcc tccaggatga cgaggaggtc gaaggatggg 1560
cagaggctgt ggagcccgag gagatgctgg atgttaaaat ggaggacatc aaggagtcca 1620
ttgagaccat gcggacccgg ctggagcgca gcatccagat gctcacgcta ctgcgggcct 1680
tcgaggcacg tgaccgcaac ctgcaagagg gcaacttgga gcgggtcaac ttctggtcag 1740
ctgtcaacgt ggcggtgctg ctgctggtgg ctgtgctgca ggtctgcacg ctcaagcgct 1800
tcttccagga caagcgcccg gtgcccacgt agcccctgcc atggaaggaa gaacgggaca 1860
aaggaggggc agcagggtgt gtgcatatga gacttggggg tccctcccca attttagttt 1920
cctgccaaaa cgggagtgtg cagtcagggc ctgcggtctg gccccatgag tctccttccg 1980
tcctcagcgg gcagggaaca cctctggctt gtagaaggga cggctcagtg gctgcaccga 2040
cggtcctgga aatctcacat ggtgggcact gcagcgttgg aacgtgagcc tcggatttcc 2100
tggcccctct actgtaaatg tgccttagcc taagcctccc atcctgtgtt agcgttgcct 2160
ggtgtggggc agggcctaac aaggaaacct gggccctcca agccaggttg aggtctggta 2220
acagaatgcc aggaaggggg cctggaagac cacctgcccc ggcccctctc ctgcaggggc 2280
cccacacagg catgagggat ggcccggcca aagtctaggc agaagcctcc tataacaaag 2340
ggtggtgtgg cctgggcatt ggagtcaata aattatggtc agaaaatcac aagagcctgc 2400
tgattggggc tcagggacac cagccaggac agaccggccc ccgagggggc agccagacag 2460
gtgggagggg cctcccacga agctcagaag aaaggtcgtg cgggagccga tggaggtggt 2520
cgcctgggca ctcctggggg aatccctggg gggatccgaa ctggccgaga tgggatctga 2580
ggcggggctg ggaagaggaa ttccaccaca tggactagtg tnctcgatnn anggnnntca 2640
actggctgtc cct 2653

<210> 191
<211> 793
<212> DNA
<213> Homo sapiens

<400> 191
taaagagtag gaaaatctct tcagagttgt gactttccta ctctttatta aaatagtacc 60
ttgaaagata gcaaaggaat gaacaactca cccttggctc ttttttcttg ggaaggctgg 120
aagaagttcc tagtccttct ccctgccttt tgtatcacac catctcaaag tacttctttc 180
tctaacattg tgcccactac ctaccaatat tgcacaccag gcagttgtca ggccgtgcat 240
tcaaacgcag tgggaggaaa tacctggaaa tgatgtgtag tgaggcccag tgtctctcat 300
cctctcttag ctaaacagct accgtgggcc tgctgaggat tgctgccttt aaggtggttt 360
actgagcttc ctggagtact gttattgtgt ctccaatagg tgagagtcct gagatctact 420
ggggggtaaa atggaaatca tggcagttta ttattccatt catcatatca tgcttggtag 480
catgctggtt tatagccaca tttctccaca gtagaaccat tttagtgtgc aaattctaca 540
tttttagggt aatttacaca atatatgatg taaagataaa cgagtctatg tttctcacta 600
accaacagga attgtcccca gtagataatg agtaatcaat caagatgaga cagaagtcct 660
ccccaatgaa aattaaatat cttatttgca aagatgcgaa ggaactaatg agccagtgct 720
actaacaaac gctgagtttc attcagcttc ctaagacatg tttaaagagt atggacaaat 780
cctcatgcta taa 793

<210> 192
<211> 2311
<212> DNA
<213> Homo sapiens

<400> 192
gcccagcctt tttttccttt ctttctttct ttctttcttt ttttttttg agctggagtt 60
tcagtctgtc acccaggctg gagtccaatg gcacgatctc ggctcactgc aacctccacc 120
tctcaggttc aagcaattct cctgcctcag cctcccaagt agctgggatt ataggcgtgg 180
gctaccacgc ctggccgatt ttttgtattt ttagtagaga cggggtttca ccatattggc 240

caggctggcc tcgaactcct gacctcaagt gatccaccca cctcaacatc ccaaagtgct 300
gggattacag gcgtgagcca ctgtatccgg ccacctttc aaaaaatatt gttattacta 360
ttttctgtct ccaggcttca acagtttaat ctattattta cagctcctca agtgtcacca 420
acaaaaatat actcttgttt tcaatttctc ttggcactca gagatgaatc tgtgtgagca 480
cggagagcta gagatctcag agcagccagg aactgagatg caaatcattt caccttgtg 540
tggtctgctc attctgtagt ctttgtctta cccttaaagt gggaaactct ttgacaggga 600
aatgccatgt ttgaactgaa gaagcctact ttatagtttt ttctctcttg cctaatgttt 660
tcagtagcgt actatgttgg gtgttttagt atgccacatg gtctctattc tctgaagctg 720
tatgatgaaa caagtaccta actttaacag gagataatta atgtcttaag ggaagtacta 780
acagaattgc tttagagtct gggatttttt ttttattgca agaactattt ttcatggtgg 840
aaaggagaat ttatattaaa gaaccagtgg gctcctgggc ccaccatggc acatttagcc 900
ttcactggga agtagaacta aatgaaagcc atcagcagcc aacacagctt cacttcagct 960
cattgtatct gaatccagtg ttcccaagga catcagtgac tttatttttt tgaaaactgt 1020
aatgctttaa aacttacttt attggatctc tttgcagctt ttgacacagt gaaccacttt 1080
cctttcctga aatgctttcc tctcttggct ttctgatgcc atgttctcct gtttcttctc 1140
tacttctctg gccacttctg tctccttaga agctcagtct tgctttgcct ggcccttgat 1200
tgttagtttt ccccagggct cacttcttag ccccttttc ctcatgtctt ataatttgag 1260
tcatctcatc tactctgggg agttgaatgg tcgcttgtat gctgaaaact cccaaatttg 1320
tatctgtagc ccagcttttc ctactaagct ctacttacat atctttgcag acctcattac 1380
ttcttgacta acttactggc tccttaaata tctaaaattg aatccatgtt ctacctcccc 1440
atacaaaaag ttcctcctcc ttccctacca cagttaacag ttaccattac tctggccatt 1500
actatcctgg tctctctgcc tcctgatggg agtcagctcc tcctgtcaag ttggcctcca 1560
cagtggcttt cacatgtcta ccttccactc tgctaatttc tgagcagtta ccacattctt 1620
gctcaacagt aggtaccagt cacctggatt gcttattaaa tagttgaatt atggagcccc 1680
aggaatgtag aaaatttact aagctaccca ggtgattcta atttgtatgg ccttagcaag 1740
taccacactg cactgtaagc tccattagca aaggaagggt tatcctgcaa gtgcagtatc 1800
taaggattat cattggttaa cttataggaa attgtcttca tctgaccatt tttgatcaat 1860
aaggtaatct taacaattca gtcaaataac tgctcaatta aatatctatc aaatgaatga 1920
atgactccta ttccacacac ttcatgataa cagggctctg ctgtttcagg ctcaatttat 1980
ttaccttgtc acctggccca taagtaagaa gactgtcatc catggccatt gttttctagt 2040
tcttctgagg acttttctcc tctctttgtc agtttaaagc ctccctaatc aaccaagttt 2100
gtttctaggc acatgcactc ttccctcctc tcatgtgctc agcccttggg taagagtcca 2160
tccttctgtg gtctcattag aaatttcaca ttcttgcccg ggcacggtgg gttcacacct 2220
ggtaatcccc aacacttttg ggaagcccag gccggccata atggagggtc gggaaaatca 2280
agaaacttcc tgggcttaca ccgggtgaaa a 2311

<210> 193
<211> 832
<212> DNA
<213> Homo sapiens

<400> 193
aacttgttta catgtgtttg gcccatctgt ttaaacttct tgtttatttt aataggagca 60
actcttgggt gcaagctcct tttgtgcttg agaccactac tggcctcttc tccagctctg 120
tttccctaat ctgcattcta aatctcttct gtaaacaaaa cttgaataac aacttttttat 180
gataaatgct gaacacctcc aagatgactg ctgcaaaata aactactttt acaaccagaa 240
agcattctaa acacacagaa aatacttttt taatgctgtc cttttcttaa aaaaacaaaa 300
agcagaaaca aaaacacaca caaagaattt tattttctgt gtgtgtgtgt ctacatatat 360
atgttcttga cgatccttct gaagtgtgtg cattcccatc atgctccagt ggttttccag 420
ccccgtaagg ccctcacttt tgattgtgga gatgatgttg aagtgcacag ggtatttact 480
gacccgccta tcttgttctg tgacaggcag aatcaatcaa tcaagaggcc agctttgcac 540
ttacatattc caggcacatc tcccgcactt caggatacga gatcacagtg ttatccagcc 600
aaatttgatc aacaaatgca gttcccctgt gccacagcaa tctcagatga agatggtatt 660
atgtgggact gagcaaagtg ataaagacag tacagcagag cctggagtaa gcccggctgg 720
gcccgctctg tcagctttta tatacgaagg cgggggaggc ccctttcttt cagattaaaa 780
aaggtccctt tgattggaag gaaaatcact cttggacata atacgggtgg cc 832

<210> 194
<211> 320

<212> DNA
<213> Homo sapiens

<400> 194
tttcgttgga gggagccgct ctccaaatgc ctccaccctg ctgcctcact tccactgagc 60
acccctttga cgaatttccc caggtgacag gcgtgggaat tcttcgtccc actgttactc 120
tttgatgctt cctcccctgt gttggtgttg tgttaggacg atgacatgct gtattggcac 180
aagcacgggg atggatggaa gacccccgtc cccatggagg aggatcccct gctggacaca 240
gacatgctca tgtcggaatt cagcgacacc ctcttctcca cactttcttc acaccagccg 300
gtggcctggc ccaatgtatt 320

<210> 195
<211> 655
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(655)
<223> n = a,t,c or g

<400> 195
cgaccatgac tgcgaggttt cttatttgtc tctttcaaac caccatgtat gcagaattca 60
atttggggca aagaagatgg caaacaagga atgcacccaa cctttctggc tggttggggc 120
tggcaggggc tgctccctgg caggggagaa tcagccccat gctagggacc aaggtatccc 180
tgtgtaactt gtctgaagag tccttagctc cattggccaa gcatactccc agggcctagc 240
ctggggcatg gaacactgtg ggtgcctaat caatgtttgt tgtgtgaatg aatgactaaa 300
aggagagcag aagtgttgcc tcagttcggg agctgggggt gcacctctcc cccagggctc 360
ctgggtggca gtcacttaaa tcttgtctgg ccccgtcctg cccagcctcc cctaggcttc 420
agcagggctg aatggcagtg taatgggaac aggcggagag ctggcgctaa caattggaac 480
atctccggtc agccggggcc cattgtggnc tcttcctgcc tctgctgccc gctagctttg 540
ctactctttg tgttgctgcc tcccaggtgg ccagcgtcat gcccatagtg tgtctttctg 600
gcctgctctg gggtctcaga aatggcctgg aaatgttttg ctgattatta ggctc 655

<210> 196
<211> 1298
<212> DNA
<213> Homo sapiens

<400> 196
gtcgacccac gcgtccgggc aatttttga ccatctccac tggaagacag agaatggctt 60
cagggccggg cacttgtgtc ttcctgagtt tcatgttccc caatcctacc attggtatgt 120
agtaggtact cagagctttt ttactgaatg agtcaactgg tgagtaataa agaaagcagg 180
aaggagtggc tgagtgaaga aatctggaaa gaaggaaggg agggaaaaca ccaggtagga 240
aatgagtgaa cgatggaaga ggagagggaa ttcaacaagg acctcctctt tggcttcagg 300
agctggggac cctgaaccag acctgtggat catccagccc caggaattgg tgttggggac 360
cactggagac actgtctttc tgaactgcac agtgcttgga gacggtcccc ctggacccat 420
caggtggttc cagggagctg gtctgagccg ggaggccatt tacaactttg gaggcatctc 480
ccaccccaag gcgacagcgg tgcaggcctc caacaatgac ttcagcattc ttctgcaaaa 540
cgtctccagt gaggatgcag gcacctatta ctgtgtaaag tttcagagga aacccaacag 600
gcaatacctg tctggacagg gcaccagcct gaaagtgaaa gcaaaatcta cctcttccaa 660
agaggcagaa ttcaccagtg aacctgcaac tgagatgtct ccaacaggcc tcctggttgt 720
gttcgcacct gtggtcctgg ggctgaaggc aattaccttg gctgcactcc tactggccct 780
ggctacctct cggaggagcc ctgggcaaga agatgtcaag accacaggcc cagcaggagc 840
catgaacacc ttagcatgga gcaagggtca agagtgaggg gtcagcccca gagtgaggac 900
cctctgagtt ggagaggagc cagggctcct caaccatttc cctacctcca gtcccagcct 960
ctaggtgccc ccaggcctca tgacaaactc ctagatccct acatctggtt ttggtccacc 1020

tagtgaaatt cccttctttg caccgggctt ccctctaaaa tgtctccctt tctcttttg 1080
gcctgttcaa gacctccttg cttttcagtc cctggctcag tctctcctca acacccttgc 1140
ccctgctgca gccctttctg gtgcgccctg ccccttccc cacctcgcta catccttctt 1200
ggcctccaac atccaactca gagtcttctt cccaggagat gtctgtaaga atctctgaac 1260
tcaaccagcc agaccatctg tgcccctcca tttacacc 1298

<210> 197
<211> 1298
<212> DNA
<213> Homo sapiens

<400> 197
gtcgacccac gcgtccgggc aatttttga ccatctccac tggaagacag agaatggctt 60
cagggccggg cacttgtgtc ttcctgagtt tcatgttccc caatcctacc attggtatgt 120
agtaggtact cagagctttt ttactgaatg agtcaactgg tgagtaataa agaaagcagg 180
aaggagtggc tgagtgaaga aatctggaaa gaaggaaggg agggaaaaca ccaggtagga 240
aatgagtgaa cgatggaaga ggagagggaa ttcaacaagg acctcctctt tggcttcagg 300
agctggggac cctgaaccag acctgtggat catccagccc caggaattgg tgttggggac 360
cactggagac actgtctttc tgaactgcac agtgcttgga gacggtcccc ctggacccat 420
caggtggttc cagggagctg gtctgagccg ggaggccatt tacaactttg gaggcatctc 480
ccaccccaag gcgacagcgg tgcaggcctc caacaatgac ttcagcattc ttctgcaaaa 540
cgtctccagt gaggatgcag gcacctatta ctgtgtaaag tttcagagga aacccaacag 600
gcaatacctg tctggacagg gcaccagcct gaaagtgaaa gcaaaatcta cctcttccaa 660
agaggcagaa ttcaccagtg aacctgcaac tgagatgtct ccaacaggcc tcctggttgt 720
gttcgcacct gtggtcctgg ggctgaaggc aattaccttg gctgcactcc tactggccct 780
ggctacctct cggaggagcc ctgggcaaga agatgtcaag accacaggcc cagcaggagc 840
catgaacacc ttagcatgga gcaagggtca agagtgaggg gtcagcccca gagtgaggac 900
cctctgagtt ggagaggagc cagggctcct caaccatttc cctacctcca gtcccagcct 960
ctaggtgccc ccaggcctca tgacaaactc ctagatccct acatctggtt ttggtccacc 1020
tagtgaaatt cccttctttg caccgggctt ccctctaaaa tgtctccctt tctcttttg 1080
gcctgttcaa gacctccttg cttttcagtc cctggctcag tctctcctca acacccttgc 1140
ccctgctgca gccctttctg gtgcgccctg ccccttccc cacctcgcta catccttctt 1200
ggcctccaac atccaactca gagtcttctt cccaggagat gtctgtaaga atctctgaac 1260
tcaaccagcc agaccatctg tgcccctcca tttacacc 1298

<210> 198
<211> 566
<212> DNA
<213> Homo sapiens

<400> 198
cttacaggct taattactta cttagttaaa tatttctgaa gactattcca ttcacagtct 60
ttttcctgaa taagccactg gatgggtact tacttgacct cctaggcagt ttgaaagtgg 120
gaaataggaa ctcaggttga tgagcattat tatttctgtc attagacagg tatacttctg 180
tcttatttac caatcttgtg taattgcatt ggcttactaa taaaaactta agtctaaaca 240
atgttcttct ggcgtctact catactgtat tcttctcctg agatcactgt ttgtctccat 300
cttttcacat ctagtgggct gaaaatgcag catttgagaa gtcactgcca tctctttcgt 360
caggctctct ttctttgctt cagtgacacc acagtcatgg gtttcttctt atcttactgg 420
tggcaatttt ctgtatgaac tcctcttcta ggacccatcc ttgaaatgtt ggagagcctc 480
aacatgtggt cccaggctgg cgtttaagaa aacttccatt cctcctgttc cagtttttac 540
catctcagta agggtcgacg cggccg 566

<210> 199
<211> 842
<212> DNA
<213> Homo sapiens

<400> 199
aacagttcag gccatagagg attactaggt ctttacaggc ttattgaaga gggaggctga 60
gaatatgctg gatgaaaaat gagcccaaat ccaccatctc ctgtccctcc agttatgtgc 120
tctggtgcaa tgtggtccgt cacaagtacc atcttcatat gcaggctact cattgtaaga 180
ctcttgggca acacagcagt gagaacctct gttgtctttc taccccacaa ggcaggaagg 240
cattgggaaa agtccactag cctggtgtca ggaggctgag gttcagtcct gctctgtcat 300
gaattggcta catgactgag gaccaaggac ttgagttttc tcatctatga aacaaagata 360
ttggatacca gtcattttca aacatttctt tttccaataa gcagaaccct tttcaaatga 420
aatcttatgt atagtgcagt tgctcagagc aggtgctctc ggctgccctc tcacccttta 480
tggtagatct tgaggagcct ccacaagcat cacagcactg gtcagtactt ctgcacatga 540
ggtctgagga ccacctgcat ctgaatcacc caggtgcaga ttcctgggct ctttcccgga 600
cctcttgaat caaacttttt gggtctggga ctgaggactc tatttttaat gagcacttca 660
tgggattctc atgtacattg aagtttgagt accactaggc cagatatttc aaggttcctc 720
tcaggcttta agtacagtga ctctgtgatc ctcttcagct ttggtgtgat ggcgaggagc 780
caaagaaagg gagttcgtga cctcacctga taaaaagatt taacagcatg tattgcgcgc 840
gg 842

<210> 200
<211> 277
<212> DNA
<213> Homo sapiens

<400> 200
cggccgcgtc gacaaaacat gcatacttgc ttttttgccc atggagcact ttaaggcaca 60
attcccggca gggtgctgtc ttcctccatg tgggcagaca gtatcctggc atctctcctc 120
ttgtggcccc accagtctct acagctgtgg caccacccac atcttgctaa taagaacatg 180
ggggtgccac cccctaccac ctgcaagcct tggtccacag tggcacagaa gtttgctgat 240
tacattccat tcatgacaac atggccaccc ctaggta 277

<210> 201
<211> 597
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(597)
<223> n = a,t,c or g

<400> 201
ctccatccta gctataaagg tcttatgaga cttttgcact gtaaaacttt acacattgtg 60
ctattcactc tgctttacaa gattttgatg gatcatcaaa atctgtcaga acatgtactc 120
tgcatggttt tatatctgat tgaattagga cttgaaaatt ctgctgaaga agaatcagat 180
gaagaggcat cagtgggtgg accagaacgt tgtcatgaca gttggtttcc tggcagtaac 240
ttagtgtcaa acatgcgaca ctttataaac tatgttagag taagagttcc agagactgct 300
cctgaagtaa agagagactc acctgcaagt actagctctg ataacttggg ttctttacaa 360
aattctggta cagctcaagt tttcagttta gtagcagaac gtagaaagaa atttcaggaa 420
atcatcaatc gcagtagcag tgaagcaaat caggtggttc gtcccacaac ttcaagtaaa 480
tggtctgctc ctgggtcagc tccacagtta actacagcca tttttggaaa ttaaagaaag 540
catattgtct ttgctaatta aacttcacca canactcttc agaagtcgac gcgggcg 597

<210> 202
<211> 654
<212> DNA

<213> Homo sapiens

<400> 202
ccacctttct ccttttttcc gatgacttcc atttctatca atggcagaac cattttcttt 60
gtttctaagt cttggagtca cagtcacaca cattttcagt gatgaaaatg tttcagatgt 120
tactgacttc aagcttttgc tcattatccc atctccaatc atgccagcac atttcattcc 180
tgagcatatc gaaccatagt aagatatttc tttacttgca gcctacctgt tatctgtatc 240
taccccccact cccactcttt tcacgatcct ggcattggaa cctcagagtc cacatttgct 300
ccccctgaac ggaggtatgt tccccaaggc tgggcctctc ctgttctttc aggcagtctt 360
caggtcaatg tccttcaaaa cgtggtgctc agaactaaac acaaaactcc aggacactgg 420
agtgggacag tggatggact gcacttccaa caaaacactg tagatctccg aataaacgta 480
agatggcgct ataagcacgc acccggccta aagccctggt gaaagctcac cagcagtggg 540
gctcagtagc tctgcgaggc gggaattagt tctcttactt gacatattag gaaatgtgaa 600
atgtagacag attaaagcac ttgctcaagg tcaaacaact aggaagtggc agaa 654

<210> 203
<211> 1246
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(1246)
<223> n = a,t,c or g

<400> 203
gcgatgcggc agcggcgacc cacgcgtgcg aattcatgtt ggtctaaacc aaaggaagat 60
aaagcagtag atacatcaga tttggaagtt gcagaagatc ctatgggcct ccaaggaata 120
gatctgatca cagcagcatt gctttttgt ctaggagatt ctccaggagg gaggggtata 180
tctgatagcc gcatggctga tatttatcac attgacgttg ggactcagac tttttcactt 240
ccatctgcaa tattagctac aagtacaatg gttggggaga tagcttcagc ttcagcttgt 300
gatcatgcca atccacagct ttcaaatcca agtccgtttc agacacttgg gctggattta 360
gtattggaat gtgtcgctag gtaccaaccc aagcagcgtt caatgtttac ctttgtgtgt 420
ggacagttat ttagaaggaa agaatttttct tcccacttta agaatgtgca tggtgacatt 480
catgctggac tcaatggctg gatggaacag aggtgccctt tagcttacta tggttgtacc 540
tattctcagc gtagattttg tccatcaata caaggagcaa agattataca tgnccgccaa 600
ttgaggtcat ttggagttca gccatgtgta tctacagtat tagtggagcc tgctagaaac 660
tgtgtgttgg gattacataa tgaccatcta agtagtcttc cttttgaggt cctgcagcat 720
attgcaggct ttctcgatgg cttcagctta tgtcagctct catgtgtatc caagttaatg 780
agggatgtgt gtggcagcct gcttcagtct cgtggcatgg tccatactgc agtgggggaa 840
aaggaagtat ccagaaggaa attcatcatg gcagataaaa gaaaaggtat ggcgatttag 900
tactgcattt tgttctgtta atgaatggaa atttgctgac atcctaagca tggcagacca 960
cttgaagaaa tgcagttaca atgttgtcga gaaacgggag gaagcaatcc ctttgccatg 1020
tatgtgtgtg acacgagaac tcactaaaga aggacgttca ctacgctcag ttttaaaacc 1080
tgtactttaa aagttgtaat attactagca catatatgca agcacctagt ataattcctt 1140
tgtaatatgt gaaactttat taatgtatta aatattacaa ctagctaaat ttattgtcac 1200
tgtgtatata atgttttgaa gtgacatcta tttttataaa gtactg 1246

<210> 204
<211> 1153
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(1153)
<223> n = a,t,c or g

<400> 204
cccacgcgtc cgtaagggtt ttgttttagt agcaggaagt tgcttgaagg aaggcagggc 60
atggtggctc atgactgtaa tcccagctac tcaggaggct gaggcatgag aattgcttga 120
acctgggagg tggaggttgc ggtgagctga gatcacgcca ctgcactcca gcctgggtga 180
catggcaaga ctctgtctcc aaaaaaagta atgggatact ttagttgata acttttttaa 240
aagaattgct ttgctaaaag ggttcatttt atgtacagaa attgtatgtt ttaagtctga 300
tgaattttaa ttatacttca cacctactag aagcaagtca caagtcagta tataattata 360
ttttgctttg tttgcagagt acgttagtat atatccaagt actttatagt agattgataa 420
ttctgtgaag gaaatgaagg cgtgagtatc agggtaacca aatgaaagtt tcctgactaa 480
aggtagctgg tttaaaccta caaacattta catgtgattg gtatgatggg tactgaatga 540
attgggttgg atcagtataa gttgatcatg tcatgtagaa agttattaaa gtagagaaaa 600
gagtttttac tccagactgg tttagggcag gtctttctga tttaaaaaca gtgtccagga 660
agctgcacag atgagcagta ctaggtcatg ctggtggctg agctggagtt tggacctgac 720
cctcaaacac aggtgcatgt gcacatttcg tggtctcctt actgggcttt tgacttttcc 780
cttatttttct cctgtgcttt atttctgtaa caaatttccc aataaaacca acatgttcct 840
tttatgtttt tgtaaaaatt atttcttaag tactgttttt tttatttttt tgagacagag 900
tttcgttctt gttgcccaga ctggagtaca aggggtgtga tcttggctca ctacaacctc 960
catcttacgg gttcaagcga ttctcctgcc tcagcctcct gagtagctgg gattacaggt 1020
gtccaccaac acacccagct aatttttgta tttttagtgt agagcgggtt caccatgttg 1080
gctaggctgg tctcgaactc ctgacctcag tgatccaccc acctcgcctn ccaaatgttg 1140
ggatacaggc tgc 1153

<210> 205
<211> 657
<212> DNA
<213> Homo sapiens

<400> 205
gaaatgcctc tactgcatgg tcacctttt ctccaagaaa gaaaagattt ctctctagct 60
ttaataagca gagctttgca tactaaataa cttggatttt cccatcatcc caacattcaa 120
accaagcctc agaggaggaa aaatgcttca gagaggctgc aatgactggg ttttgccatg 180
tgtttctgag aatcagaaag gcactgcaga tcttctagtc ttctggagtc ctgtggtctc 240
tgctgagcac tgccactggt gccctgattc aaggcccaag cttgggaaac attgtttctg 300
gagctgctaa accacaaccc agaaaggagc acaagggagt gatgaagggc atcctctttt 360
ttttttttg gaaaggggtt tatttttccc caagtttaaa gcccaggggg gaaatttggg 420
ttaattgccc ccaaccttgg ggggaagggg ggcccattgg gggaaaaatt aaaaatgggg 480
gggttttttc tggaagggaa tttttcccca ccatggaaaa gaaaaaattt cccccaaggg 540
ccaaaaccaa aattaaccct cccagaaaaa tgggggccca aaggaggccc accccaaaat 600
ggcccacccg gcagggtccc tttaatagaa gcccaaaaaa gggcaaaaga tatccgg 657

<210> 206
<211> 365
<212> DNA
<213> Homo sapiens

<400> 206
aaagtgtctg gggcggctgc tgaccctgtg gctgggttct tcatcaacag gagccagagg 60
acagggtctg gacggctcat gccttgggtg ctgggctgca ccccccttcat cgccctggcc 120
tacttcttcc tgtggttcct gccccccttc accagcctgc gaggcctctg gtacacgact 180
ttctactgcc tgttccaggc cctggccacg gtgccctaca cagcgctcac catgctgctg 240
actccctgcc caagggagcg ggactcggcc accgcctacc ggatgactgt ggagatggcg 300
ggaacactga tgggggccac tgtccacggg ctcatcgtgt ccggcgccca cagaccccac 360
aggcc 365

<210> 207

<211> 649
<212> DNA
<213> Homo sapiens

<400> 207
aagttgctta tttgtagcct ccttttgagg ctacaaggct agtccatatc aaagcattca 60
ttcatcctgt agaccgatag gtatggaggg gaagtattat tggagagacc tcagggtttt 120
ctcacctgag aagggagaga gccctggggt ggtcagagct gggccctgag tggtgggatt 180
tagaaagaag ggaagagggg agggcttgga agataggaga aacagcatga acagcagcag 240
ggggaatggg tgagtgtcct cctgaaagga ctgtggagat ggccgtggag cctcttcttg 300
cccacttcct gagatggtca tggcttagtg ctagagactt ttattccttg ggcaatgtag 360
acccagctct atgggttccc tgtttttttcc ttctcttcct cctcattata acagataata 420
ataatgattc ctagctttca gggaacacat cttgtatacc atgttctgtg tgctttacat 480
ggattgtctc attacagctt cacaacaacc ttagagatag atgttctgta tcattatgtg 540
caccttacag ttggtgcagc tgaggcgtag agaggttaaa taacttgccc ttggtgggtg 600
acatagcaag gcaggagttg ggtgaggagt ctaccagggc agcctgaaa 649

<210> 208
<211> 369
<212> DNA
<213> Homo sapiens

<400> 208
ttacagtctg tctctgaaga ttacttgtac agtcccatca gctatacttt ttcaccccac 60
agtccagctg ccagcctcgc cagaatcccc ttattatttg aaatttattg ttgtattcac 120
tcaatgtatt ctgtgcctga aatgcatctc tggagtcaac atgggtaccc tttccttctt 180
taaagccctt ccctgatgtg cagtgttacc tgtggtgtcc tcttcgctct ctcaggcctc 240
ctgttatact cttctccttc cccacactgg aacagaccca gtagaattgc agtgtatctc 300
atgtgcctca ctaagtactg cactgggagc tctgccgcct cctgccagtg actctgcttc 360
agccatgtt 369

<210> 209
<211> 631
<212> DNA
<213> Homo sapiens

<400> 209
aaatcagtga ctatcctggg atactttatt atgaagggaa agaaactaat attgccaggt 60
cctaacttgt tcggtcctgg actaactatg tgggatgcca caggtcaggc agtctttgaa 120
gtaaccttat gaagtaggaa ttatatcgat tttatataag agaaaactga agccaacgag 180
gtaaaataac tcgtcccaag taatgaagct tgcctaggtc ccagccaggc ttagtcaagc 240
cctggtcctt ctgatttcaa agcccttgct gcttctctgt attgcacact ggttctctct 300
gttgcacact gagcaacatt gtctttttaa aatctattat gaaataaagt taaacagctt 360
cataattatc attgtcatta ttttttctttt tagctactat tttttggtct gtcctgtgga 420
aacagcttgc aatcctcaca agaacctcgt aatgtgaata tgccagcgtg ttgctacaga 480
ccgtgcttac ttcagccaat ttctctgctg aatatcctcc ttttactgat gaggaaacca 540
agtcaagaag tcataaatga cacacccaag gctgggaagt ggttgagcag gtatttagac 600
tcaggtctat tttactcctg tgcctgtggg c 631

<210> 210
<211> 818
<212> DNA
<213> Homo sapiens

<400> 210
gcaacctggt tggatttaag gtggaggtct tcctctctcc aggtggctgc cctcatctct 60
acccccaccc caatcccatg cattctcctg ccctcctgct ttgggcaaga cccgggcgat 120
gtggagttgg catgtccagt tgcaggtatc agcaccactg caccatctcc tctgccttca 180
cttcccacca gcacacagaa tttatatgcc cttcccaagt ccaaaaaggg cccctgccat 240
gctaaacaaa ggtattcata tgcaaggcat gagctctgtc tcatggaagg gagaggccaa 300
gttctctttc caccatcaga gagtagcgtt caacatcatt tatacaagac aggcttttgc 360
ccttttagtc ttgctgaact agaatgaggg atttgtgtag tcgaaagtca agctgatcat 420
tgtcaagttt cgtggaagtg gaacctgaag gaaggttagg tggagactgc agagagggtt 480
agcgtcctgc cttaccacaa ggtaagtcta gaagagggtg gggaagggag ttcagaattc 540
ttcattcatg gagtgacctg ccttggtcat ccttttatct tccagtcaga ggggtcatca 600
gtcaagagct ggccaggctg ggtgtggggg ctcatgtctg taatcctagc actttgggag 660
ggcaaggcgg gtggatcatc tgaggtcagg agttcgagac cagcctgacc aatctggtga 720
aaccctgtct ctattaaaaa tacaaaaatc agctgggggt gatggcgggc acctgtaatc 780
ccagctactt gggatgctga gacaagagaa ttgctttg 818

<210> 211
<211> 387
<212> DNA
<213> Homo sapiens

<400> 211
gatgattgct gtagggtttc accacgggca actctaactt gtcacagtca atccacacct 60
tgtcagcatt taaccgcctc acacagaatg caggtatctg gcccgaggcc acagctgttc 120
ctgccctctg tgttctttgt gctgttattc tcatatactt tcacagaaac cacacaatgg 180
actgttgtga ttttggcttt aaacagtaag ttatctttta aagagattga aacaattttt 240
tagtctttta gtcttttatc cttccctcta ccatttctga tgggcttcat tcctttgtgg 300
ggacctgagt ttccacccgg catcatcttt ctcctgtgtg aataacttcc tttagcagtt 360
gttatagaac agatgtcgac gcggccg 387

<210> 212
<211> 400
<212> DNA
<213> Homo sapiens

<400> 212
tttcgttttg cctttacata gttatcagct aaacctttga attaaaaaaa caaacagtat 60
ctaagcttta ttgagcaatg tttatgccag gaactgtgct aaggatttta ctagcattac 120
cttatttaat ccttacaaag caggtacaat tttttctatt ttctgatgaa attatggcat 180
ggaaagtggt ggcaccgggg cttgaactca gtgctgtgac tccagactcc acactcttta 240
accattacac tatactatcc tgacattaat attctatatg attaagcttg cggagaagta 300
aattgaggtt tgatatgact ttacaccgtt ttaatcaata ctgtgggttt atggaaaggc 360
acagtgttat caaggtgtat aatattgggg agggtatgga 400

<210> 213
<211> 567
<212> DNA
<213> Homo sapiens

<400> 213
gtcagtcaac aaatacctac ccaagataaa agaggaatta cttggatgtc caacatgcaa 60
tttaaaaaac aagataaccc aaacatagct gttttggctg ccgtcatctt agcacacaaa 120
attttgaaaa cattttaatt catgcaagtc aagttcattc ttaaatatta catttcattc 180
ctctggaaaa ctgtaacagc caatggagaa acagtaaaca tgtctcttct ctacattttc 240
actactatgg aaatgagaaa gaaaagtgaa gtaggactgc acttacccat ttccattctc 300

aaacccttct ttacaatagt gctagatgaa aagattgtta caggccaggt gtggggtggg 360
gagcttttcc tgttgttttg caaagattaa ttttcagaac taaaactggg ttttcaattt 420
tgctaaactg aatgagcctt tttaagaagc tgtcctctgc tctaaaactt aatttacaag 480
gtatcatgac atattcattt ctgagaaatg cagtaacacg cagcagaagg aaccagttaa 540
atttcttcta tcatgtcgac gcggccg 567

<210> 214
<211> 806
<212> DNA
<213> Homo sapiens

<400> 214
cccgttctga cgatttcgtc ctagcgccca tctttatgtg gggctcccta gaaacccttc 60
tcttttattt gccacttcct gttttgtgga cactccatgc ctgagctgcc aacgtgggtg 120
ctggcattac tgcctcatcc tgtggttttg ctcattgact ctggagaact agaggccttt 180
gagcagatat gcagaagcac cttgaaggca gtgtggcact cagtacatgg tgccatgtct 240
gtgtgcttca tctgcttcac cttttgccat taaagcagtt tccaggggac tgtattccag 300
catagtgttg ttttaaaatc ttcttagtta aggtgatttt tacaacctct ttgctttcct 360
tcattgtcca gcaagtcttt accatgagaa atatcttatt gcagtctagc ccaaatgttg 420
tacttttatt tattcttgtt tcttttaaca tgccatagaa gagtttaaag ctgtctgtta 480
ggagatagaa ttggggccgg gcgcagtggc tcacgcctgt aatcccagca cttagggcgg 540
ttgaggtggg aggatcggct gaggtcagga gtttgagacc agcctgacca acgtggtggg 600
acccccttctc tactgaaagt acaaaaatta gccgggcatg gtggcgggcg cctgtggtcc 660
cggctcttcg ggaggcctga ggcaggagaa atgcttgaac ccgggagtcg gcaggttgca 720
gtcagctgag gttgcgccac tgcactgtag cctgggcaac agagtgagac tccatctctg 780
gagacacaca cacacacacc tgtatt 806

<210> 215
<211> 459
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(459)
<223> n = a,t,c or g

<400> 215
cctgaattcc tcgactttcc cttttgggct cacatgagaa ttgttaggag aatgtgcatg 60
tggagtgcgg gcccagcacc tgctaccgtc tgcgctgtaa tggtggctgc tcctaagagt 120
cctcagtctc ctccccgttg ggcttgtgtg tacagcctta tcggttgtca ttcttcggat 180
ccattctcag tttatttttc tggtattagc tggagggaca tctccttaag cctgtactct 240
atggctcagg agtctcaaaa ccagtccatt ttgaagtgag agtatcccta ataaaaaggt 300
gagtgtcccc actcctgtgc cttgtttttt ttgtttgnnn nnnnnnnnga aacggagtct 360
cgttttgtcg cccatgccgg aagtgcggaa ggcagtggta aagggaggtg gcggcagcgg 420
ttagcggact caagtctaaa ccgggcgtcg acgcggccg 459

<210> 216
<211> 881
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(881)
<223> n = a,t,c or g

<400> 216
atgaagctgt gtgaaccaaa tcctctttag gattagagaa gggtacattt ccttatggga 60
aggtgagaga ctgaggagcc ctcattctcc cttactgtga ctgcattgtg gagcaaacct 120
cagttcaggt acctgcttag gccctatata atttcttgat gtgcccattg ctggtataca 180
agatcatact tgtttttgct gccatgtttt tcttctcaca gggaagccaa gttgagatca 240
gatcccatga gggtgaacac tgtgtgggaa cagtacatct tcttagtcat tttctgtatt 300
ctaaaaataa cccagtattc tataagggaa ataccagttt tatatttgaa acaatggagg 360
aggattcact ttcttctttg gctgagcgca gtggctcgtg tatgtaatcc cagccctttt 420
ggaggctgag gcaggagggg tgcttgagtc caggagtttg agaccagcct ggcaacataa 480
ggagactgtt gctatggatt accaaaaaaa aaaggtaaaa attaaaaact ttaaaaaaag 540
atttactttt cttcccttaa gncccctata aaggatccta gcaggattga tgaaaatgga 600
taacaaattt aaaaaattgg cccaattagg cctttgaaat tatttttgga aagaagaaaa 660
tgtccggcgc cggccggcga atttccagaa actatggaat tggtaaggta cttggaaaaa 720
cccccggcaa ggtccacctt acaacttggg ccatcggggg caccaaaatt aaaatttctt 780
ttcaatttaa accatgcggg gttgcccttt ccttttaatg caacctaata cctcccccta 840
aaaggtcaaa atctttgggc caagggaaac cctcttggat a 881

<210> 217
<211> 700
<212> DNA
<213> Homo sapiens

<400> 217
cccacgcgtc cgcggacgcg tgggtttgaa catgaaggca gacctgaagc gagtgaggga 60
gtgagctatg cagctaacca gggaagagtt ccaggtagtg gggatagcaa gtgcaaaggc 120
actgaggaaa gggcacatgg gaggggggcag caagaaggcc actgtagcca gagcagtgag 180
tgagggaaaa ggagaaataa tttcagagat agcgagtgat cagcatgtat agggtattaa 240
gcataagaag gactttggat tttatcctgt gtgagatggg aaaccattgg agtgccagag 300
taatttttc ttaaagctag tctgcttgtc atttcttttc ttaaaaatgc ctcagttccc 360
tgttgccttt ggaataatgt tcacgtactt tactttggca cacaaggtcc ttcacagtca 420
ggcctctgcc tgtttattta gcatcatctg cttctttccc acatgtaccc ttcactttag 480
ccaagtggga tcacatgcag ctccctggat gggtcatgat gctttgtgtc ttcgcgtctt 540
tctgtaccgt ttgccctgtg agaagccctc cccctcagct cacatggtca ctggctctgt 600
tttagagggt cctctttgtg ctcttgctct ctcatcattt ccacccggag ccactttgca 660
cttaagctgc ttaagcttga aacgtgctgt gttcttttac 700

<210> 218
<211> 792
<212> DNA
<213> Homo sapiens

<400> 218
ttcgttgtcc tataaaatcc caaaatgagt gttgcttcat gacagtttct ctttctgttt 60
ttattgatga gttaatcttt ttctccgatg tgggcaggga tttagacata gtggactagt 120
ggctgttgcc tgtttctagt gaggacttct gctgcccttt ctgttttgag ttcccattct 180
gcaaaccccc cctcacatgc ggcccctgcc gttccctacc aaatatctag gtaagttcag 240
ttggaagtgg ttttttcttt ctctgaatgc tggtgactat catagagaat ttagcacagt 300
ctactttgaa tttgaaagct tcgggcattg gtcttttctt cagtgcttca tgtcctggag 360
aacaagaagc atgcacacac acatttcagt ttcttttaaa ggaaaaataa gacccacatc 420
tgcttatctg cttctttttc ttttttttctt ttgttatgga gtctcgcttt gttgcccagg 480
ctggagtgaa gtggtggcgc gatctcggca cttggcaagc tccgcctccc gggttcacgc 540
cattctcctg cctcagaatc cggagtagct gggactacag gcgcccgcca ccacgctcgg 600
ataatttttt gtattttag tagagacggg gtttcaccgt gttagccagg atggtctcga 660
tctcctgacc tcatgatcgc ccgcctcggc ctccccaagt gccaggatta caggcgtgag 720
ccaccgtgcc cagccactta actgtttttt aaacaaacat gaacttaaaa ttatatcaaa 780
cagaccatgt at 792

<210> 219
<211> 375
<212> DNA
<213> Homo sapiens

<400> 219
gcccatcaga ggcattgttc ctttctgtta cattgttttt gattgctacc atttcctttt 60
tattcttagg atttccgtct ctctgctgac attgcccatc tcttcttaca tgctgtctac 120
tttatccatt ggaactctta gcatgctaat catagttgtt tcggattcct ggtcttacag 180
ttccaacagt cctgctatgt ttggttctga tgctggcttt atcccttcaa attgtatttt 240
tgccttttag tatggcttgt aatttttctt aataactgga catgaagtac ttattaaaag 300
taatggcata aataggtcct taggaattta ctggtaaagt gtaggaggag aggcagtgtt 360
ctacaggcct atgat 375

<210> 220
<211> 1128
<212> DNA
<213> Homo sapiens

<400> 220
tttcgtggct tgatgaatta ccacacagtg aacacacctg tgtacccacc acctggatct 60
gacacccaga cgctggcagg cctcccccaa ttaccatgcg cttcatctcc ccaaaggcgc 120
cctgtctgca tttctccacg gatcatacgc ctgtcggccc ttgtgcctgg cttcacgcta 180
tatcacttct gagagtcatt cgtgttgcct gtagcagtgg gctgttcttt ttatgctgtg 240
caagtgtcaa ccgtaggaac ctgccaccat ttatgcccat gctctggggc tgctccccgc 300
ctggggctgt catccacagc ctttacaagc attcctggtg aagcagagtc acaggcatct 360
ggcccaatct gttcaaatca ttctaaatgt tggcatctgt tacgttctgg tgccccaggc 420
cacacctgga tcaccccgaa ctccctgcgg aagtacctgt tgctgcctcc tgagagtcct 480
ctagcttaag agtcttggtt ttccgaccag accctcagtc ccctaacccc agagtggttc 540
ccaggacccc agaccaccag ctctcctgtc ttccaggcta gactcgacct ctcatgctca 600
gagggtcatg tggacctctg gcctatggac ccaccctgtc cctggcttca ccctgcagcg 660
tggccattgc agacaccact tgccttgccc ctcttgggta caggctcttc ccccatgccc 720
atcttcaggt ggagacctcc tgtgcatctt ctaagcatgg ctcagggccc ctccttttg 780
gctggggcgg ccagaggaga caaggccaag ggtgcaccga ggaggcatgg tgctaacttt 840
gccctaacac gctgggccta cccaatcagg gccttgaatt tgctgggggg caggcaaaca 900
tggtgacgtt caagattccc actttgcttg ttgggcctcg ctcaccagcc tcttggcaca 960
tgatgggcgc ttaaaaacaa tttaagcggg gggagcaccc tcttgggccc aactaatccg 1020
cgcgggacac cccccgcgac acgacggccc atacgttccc cgcctccgcg ctcgctgcgc 1080
gggcggccac tcgccgcttc gcccccggcc cgccccgcac gggcgggc 1128

<210> 221
<211> 650
<212> DNA
<213> Homo sapiens

<400> 221
cttatttacc ctgtgcctat catgtagcaa gcacttcata aatgttagct attattactg 60
ccaaataatt ttaactttac ttgggtcaac ttcttcacag aatgatataa aggagtatct 120
aacattctcc atccagttca aaccatgatt cttgttagcc ttttgatctt gatagtagag 180
cccctgtttg cctctctgac tcctctgagt ctctgttttg agtgtgttgt atttctaaat 240
gtaggccaac atcttactga ccaaactttc tctttaaatg ggctgttgtt tttaagcaat 300
agttgattgg gagggtgggg tgtgagttgt ggaatactta aaagccaggg ctctggactt 360
gagaaaggcc tgagtttgaa aacttgactt gcttttcact atttgatcaa aggcaagtta 420
ttctctctcc agtgttttcc tctgaaaaat gggaatagta atatcttcat agggttgctg 480

tcagaattaa gtgaagtaat gcatgtgaaa tgtttgacat gatactgaca cttcacaagg 540
gttaaataga tattagctgc tgttggttta taacattaat atccatggtg gctgagactg 600
tggcaatgtt attcattgtg tatttctagt gtagagcaag tgatcagtaa 650

<210> 222
<211> 655
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(655)
<223> n = a,t,c or g

<400> 222
tgcactatta agtagaccat gtctcccctt ctcccactgt cctataaatt ggtcctgtgt 60
ttcccaaccc caaatggggt agtaactcat ggagaacaaa atgccagtag tacagatata 120
gaacatgggt taaaaactat tctgatcaaa cccccagcaa gaatcttgaa gaggaagaca 180
gaggggggaag aaagcaatag gcttactctc cctacaactt aattccatgc tctaattcag 240
agcagtataa ttaaccccga gtttcactgc taacactgaa gtttttcagc ctgatgagaa 300
tatcatttag tatcctgcaa tcgccaataa tagcaatatg acacaatggt gcccattgag 360
gctctaatca gagcacatag tcatttacat agaacatttt ctgtgactct ctgtaataca 420
acgccacaaa gactcaattg agacacaagt ctccctttcc cataanctan nnaaaaaaaa 480
aaacccagaa aaatgggttg gggggggggaa aaaggaaaag aaggaccccc ctttttttttt 540
tcaaggagaa ccaaattgtt ccccaccatt attccgggggg gtttacaggc ccggggtttt 600
tctttaaaaa aggttatttc cggtgcctga aaaaaaaatt aaaccttgtc ctttt 655

<210> 223
<211> 665
<212> DNA
<213> Homo sapiens

<400> 223
tttcgtcctc agcctccttg cagcaggagc cacagcatcc cactgggacc cttaagcagc 60
ctgggcacag gatggtgatg ggtctgtcct cacttgggcc ctgttgccca tgacccacct 120
gtcccagctg tgtgtgtcag tgcgtctgcc tctctgccag atgttggagt ctcccagtat 180
gcccactgac tatctttcag aattccacag gggcagatgc tcttaatgcc ccacctgagt 240
ccccttggca tttcttttct gtttcaggac acatgggccc tgtgtctggc tgctggcaca 300
tgagcctctg cttgagggtt tatttggcac tggaccctgc tcaccaggaa ttgatgcctc 360
ctgggagtag cctccaacca attactctag gaattggcat tgagatcctc cagcctccta 420
cccttgaagt gggcaactct gaggcactct cggtgccgtc tcgcagaact cccaggagga 480
ctgagctccc gtggcccacg gtgctaactg gctttctaat aaacacttta tgagccttct 540
tcccttccct gtctcacttc cctctccctc ctccccatat cagtgcttcc tggggtcaca 600
tctcaaataa gattggatcc ttggctcaag gtttgcttct tagggaaccc aaaccgacgt 660
gtatt 665

<210> 224
<211> 567
<212> DNA
<213> Homo sapiens

<400> 224
ttctattata aaatgcttac gtgtgttcca gagcgactat tccagtgtca ccatttaatc 60
agaatgacat gcctgttcat gattctagaa ttcaggttgt ttaaatatga cagtaaccta 120
tgcagtcatg tgatcattaa tcacccacaa gttcagggac gccaaagatg atccagaggt 180

aggcatctgg ctatttcaac ttcacgaggc caaaatggaa gttaaagagc taatctgggc 240
tgaacagaag attctggcag aacactgaca ctgaatccaa ccttcacgca tcctctctcc 300
acatcattct ttaggcctta actctttatt ccacattctc taataaacct ctgtagctct 360
tcccacccac agagacagca tcctgagagc tctggggaac actcagatgt cctaaccaca 420
tgttaacctt gaattcccca gttccagtgg ttttcacctc tgctccacat gagataaccc 480
tttcaaagag ccatgcccag gatcttgtca tcagctgcac cacgctgcct agaaactgtg 540
agcttggacc attcactttg tggcttt 567

<210> 225
<211> 1212
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(1212)
<223> n = a,t,c or g

<400> 225
gtgttcgcag aatttccggg acgtcgactt cgttgtattg ggaaagctca gcggaggaag 60
agctcctgat gagtgagtgg tgtagctatc tgcttcatgt ggggcgctcc tgctctgcag 120
tgtatagtat tttttaggtg gactaggagc aaatgcctgc ctgatacagg caacgtttgc 180
acaaaaaccc aaaggaagaa ggctgctggt aggctagggg tggcaggtgg gattgccctt 240
gggctttaaa gaaggctgag aaaacttact tccagagtgg gaggcaggag gagtgagcct 300
tggaacaata gtggagtgtt tgtggtggac ggacgggaga gaggaacact tgagggcatg 360
aacttacttc tgtctcttta cttgaaaaag caatccactg agctaagacc ctaaaataca 420
tcaacccaga gggcttcctt tcgagttttt agtttttaga tcattcggta tcctcaatgc 480
aaagatgtga taaagtcaag ctcaggacta cttggggacc aattctcatc ctcctcatgt 540
gttacaattc cttagcagtt gctttaaaac taacattttt ctaatcttga aatatttaat 600
aaagcagctt tctgctaaag agctgaaagc attntacgaa ccttatcttg aggatatccg 660
cgcatatttt attattttaa gtcatttttt ttagcctaca gaatttcatc aacttctttc 720
cttcaggtac acggaaagca acctcccttc cccattttat agaagaataa acaaatcttc 780
ataatgttta cataatttac ccaagtcact tgggagaatc tggaatcaaa cgcaggtctg 840
ctggttccaa aaccttctcc cctgtggcat taggtatatc acatggactc ttaagtaagc 900
ctcaaaatct acaaaagtgc agtgatctgc ccagactggc cagccacaat gtcaggggcg 960
aagggtggga taacccacat gcccggagac aatggcggag acaaacaggg atcttctctc 1020
ttctgagtcc aaaacaaaac ttggggtctc cctgggggcc acccacagcg gcccaccct 1080
ttcgtgtttc accagcatgg gcggaagtgg cagcttcttg ttctcgtggc gggagcggac 1140
acagagcgta ccgacactat ctttttctaa caacggtcct tttgtggcgc gagcggcggc 1200
caacgaagcc ct 1212

<210> 226
<211> 688
<212> DNA
<213> Homo sapiens

<400> 226
tttcgtatta agataaaatt ctttcatcat gtctgtgccc tctaaaataa tactcctatt 60
gctttgtctt ccccttctcc cctcctagtc cttctttagt cattctattt agaatcaagt 120
cgctcatgag tttaagaatt agagcagcaa gaaattgggc tagagatgta caaaagcttt 180
ggacaatagt agttttgctt gtcctcattc ttattagaag tgctgttaat ttactgataa 240
attctaggac ggaagacaaa tctttgcaac tggtactata tcagtcggta attatttgtt 300
ttccttagac catattgagc tcccgaagta cagaacaatg agtagatgtg aaccaagatg 360
ctgggatttt ttaagtttca aatatgtgtg tgtatatata agagacagaa tcttgctctg 420
ttgtccaggc aggagtgcag tggtatgatc atagctccct gcagccttga actttcaggc 480
tcaagagatc ctcctgcctt agcctttcaa gtagctaggc ttacaggcat gtgccccacc 540
tggcttttt ttttttcct cttgtagaga caaggtttcc ctgtgttgcc caggctagtc 600
tcaaactcct ggccactagt gatcctcctg cgttggcctc ctgagatgct gggattacag 660

tgtgagccac cactcccagg cttgtatt 688

<210> 227
<211> 344
<212> DNA
<213> Homo sapiens

<400> 227
gctccctatg ccccaggatt aagtccaaac accatggtgt ggcatgtgag aaagtcttcc 60
tttgtctggc ttctgcagct cttcagcttc atctcttgcc actctgtcat ctctgtgtcc 120
ccagtgcatg tcccatggac acagtgtgca gtcatacccc catacaccag ctgtcctaaa 180
ctatttgcaa tacagggagg aaggttctag atagccagat ggcttcattc cgtagctttg 240
ttccttccct ctgtttggga tcctcttgac aagcacagaa ccatctttct aggctccatt 300
aaagcatacc ttttctgaaa gacccttcct ggtcgacgcg gccg 344

<210> 228
<211> 1211
<212> DNA
<213> Homo sapiens

<400> 228
tttcgtaaga tcttaattag cttttatttg agattctaaa gttgggcaac acctcattct 60
tagagaatga gggtttcaat gagctgagca gaggaagttg gttttataga caggggctga 120
ggaaagcaga aacagaaaac aaaaaatggg tggtttcaaa gttaatttat ttgtaaaggt 180
taaagcagag ggaagtccct tatgctggct aaaactggca tgtttggggg cttggctatt 240
atctcttcta atttctcaga agtcagatga acaactttgc tttggcttgc tgacatggaa 300
ctttagcaca agtgactcca ttttggtttg gtttgttggg cctagggcag ggactcagtc 360
caaaccagtg gtctcctata aattttaaca gatctattag tacttgaaag aaatgtcaaa 420
tttcatttag aagttaacaa aaatgaagat gtagctgttc agacccttta aattcatcca 480
tgttcacaga ccctctaaag tctatccatg gaccccaggc taagaactgc tggcttaggt 540
ggagcagcca ggagaactgc atccccccgg ataggggatc gtaaaggcat tagaggaggg 600
aggcacttgc cttgcagtgg ggcaactcca aaaaaggcaa agatatctcc cctaagtgtg 660
aaattttaaa aaaatcaata gctaacaaat gaagctggaa aaaccacttg agacaatggg 720
aaaatattaa tgctgaaacg tattagatgt taaagaatta ttgagttttt ttcccttaag 780
aatgtggtta acataaaacg cccttaaaat gtatgaagaa gtttccttga caagccttgg 840
attaactctt aacagattgg gggacgagga aagtcggcaa atccagatga ataacaattg 900
cttgtttggg gccgtctaat agatccaact tacgacccgc gcatcccaat ttatagtttt 960
tgttcgggcg cgtacagtca ctcccgggcc gcatttacac acccctcagc gggaactcgc 1020
tggtcttcac tcataacaga cctccaccct gatactgcga tgcgtggtgc ctcggctgac 1080
ttcttcgccg aaatacgtct accttccagc tgcgcccgct acgcccgagc ccgccgcact 1140
cctggcgcat cgcattggct tcagctcccg ccgtgtgcgt cgccgcgtgt gcgcccggta 1200
ccatagaacc g 1211

<210> 229
<211> 653
<212> DNA
<213> Homo sapiens

<400> 229
agtttccatg acattttcac tgcttttaca aacccactct tagtatgttt tcttgcatta 60
ctacattcag taatgaatat aacatactaa gagtgggttt gaaaatgtca tggaaacttc 120
ctctgctttt acaaacccac tcttagtatg ttttcttgca ttactacatt cagtaatgaa 180
tataacatat actcctccaa aaaagaaaaa cgaaaattgc tccaaacccc tgatcctgac 240
cagttcccta gggactgtac aataagttag aagcagcatt aacactacca ccggcaatta 300
aaaaacaaaa taaccaaaaa taacgaaatc agaaaatgct gcagtaatga cagtggccta 360

gcagagcgag cgagcaggct gagggttaca tacctcctgg agtttggggc aaggaaaagg 420
agagccaaaa tcatggaaac atacaaatga gggagaaaag aaaattaaaa caaagaaatt 480
aaacccgatt ccttaatctt aaagtgcaat gtatcagtca cattcatgcc tgcctgggga 540
gagagctgaa ctaaaaacac tcaacgggca aatgatccca ctactgtctg gcacgtgaac 600
gacgggctct tacatgccac gttccacctc atcaaggcaa gattgtcgac gcc 653

<210> 230
<211> 1033
<212> DNA
<213> Homo sapiens

<400> 230
tttcgtggaa ttatttcatt tctcatttct tcttaggtta tgcgatctac ttaaaattat 60
atagggtagt gtcctcaagg aagtctcctc ttccaggaaa ggaccagtgt ccactgaggg 120
accttcttca ggacatgccc caggtcagtt gttggattag gcaggccttg ggcgagtgag 180
gcttaatatg gcttgggact tgagtttttc ttgctttccc agaattcatg agcttttgtc 240
atttagaatg tttcttactc ttctgtagtt gtcctcactg gcaaatccca ccagatattc 300
ttcagtgagt atgcttcctg agtgaaatat aaaggtgaag gaacattggg cagagcagta 360
ttaactaccc atggccacat ttttgttaat tattttggta tttattcctt gtgtgggtgt 420
gtaatttttg gtaaagttgc attaaaatag tatagcatgt gcgcattttt acttcccttt 480
aaacttcttt tcttcttgga aatctctttg gctatgaaat ctcattttcc tttcaccctt 540
ctgattctta gccgagtttt acttaagaaa acgttgtatg tcttgaaatt aggttggtta 600
attacaattc cttcgaatga tttgacttct gtattcaccc tgatgataca tagacaaaat 660
caaaagcatt tttgacaaag taatgctttc tcacatattt agcttcttta tagacactta 720
cccatcttgt taaaacccaa aactagtgat tttacgttgt aaaaagcgtt tttcaaagca 780
tgtgggccat gtaggatgta cacgacagtg ggaaagacct atccctttt tgaagggaga 840
atcaagcgct cggggagcga atttggaact ggagatgggc acctccaggc gactctttct 900
ctgaacacca cgtcgcctga catcccggga ctagagagag ggctactttg ctcttggtga 960
agacctcata cgagcccacg gccctagcgg tgacgagctg ttgtccaact tgaagagcgc 1020
gctaccttat gcg 1033

<210> 231
<211> 401
<212> DNA
<213> Homo sapiens

<400> 231
tttgagtcta ctaacagtag actattagtt agggagaact tactgaatgc tcactctgtg 60
cgagactctg tcctgaatac tttttcctgt gagatagatg tgcttagtat tttcacttgg 120
cagatgagga ccctggagaa caaagggtat ggaactttcc aggcccccac aatgtgtaaa 180
gcacaagtgc aggattcaag ctgcagagtc tgagtcccaa aggaaagcat ccaggacatt 240
acctggagca gagtgggcat ggagtaagtg acagtggtga gttttggtta ttctgtctcc 300
acatttaacc aggcttcgct gtagcctaaa catcccacct tcttggtgat ttcatacaca 360
cattcacctc ttgttgaaaa gaaagcgtga gaacctaaga a 401

<210> 232
<211> 206
<212> DNA
<213> Homo sapiens

<400> 232
catctgttcc ctcctctcta tattcattga ggccaccatg tgcagggctt tattacttct 60
cctgtgctcc cctaacagct ctttccagtg gctgcctctt ccagtccatc ctcataccac 120
catcagatat aggtcctaca acatggttcc agtcaagctc acaaatgttc agtgacattt 180
agtgacacat gctgtcgacg cggccg 206

<210> 233
<211> 649
<212> DNA
<213> Homo sapiens

<400> 233
aaacattctt tccctgggaa gtttttttatt gatccttgtt tgaagttgat acataatatt 60
atatcataac tgtgtgatgc tgtttctgca tatagctgaa tgcagctttt tgagactaaa 120
ggttgcattt ccttcatctt tgaatttcca gcccctagca cagttcctgg cacatatttt 180
ggaggtattt tataaatgcc tttggaaaaa aggagttcaa gtgtttaatt tcctggcaaa 240
ctaaatagct ataaagagag agcttagaaa gtctagagta agagatgatt agcatgcttc 300
ctagaccatc tgtcactaat accaattgtt tctacttggc actgcagact tgttttttga 360
gggggagggc acgcttgtgt aaagttaaaa aaaacagaac aaaaccaaat accagaactc 420
ttaggttatt ctataccccct ctgtttagtg gctaaaaagt cattgctgat ttctaaaaaa 480
catggacatt cttattcagg agagatagat gccattattg ctttttccaa gagcaataca 540
gttgctataa ctgagagctg gttctgcctt gccccccttac aatctattct tgatgcggca 600
accccgagtg atcccattca aaggtgcaag ggcaagaaat atgtttctc 649

<210> 234
<211> 388
<212> DNA
<213> Homo sapiens

<400> 234
atcctcgact atttatccgt ctacaagatt atccttggca tactcatgta ccacaacgta 60
agcacaggcc tcctaaatat cagacatcag tcgcaaaaca gtctcatagg ttctacgtgc 120
acatgtctag taaatgcgtc tatggatatg agtatccagt tagtcacata taggacttaa 180
attcagtaaa tgcaaggcta gactcattac ttacccagca ttcgcttgta taagtatctc 240
aaaaggaagg aggtgcagat actcaagcag aatcagactc ctcaaatatg ggacatagag 300
gacatacatg aagatacagg ttatgagaca gaggacatac cattcagctg tcctatgtaa 360
gatggcatca gtatctaaag gcatgggt 388

<210> 235
<211> 313
<212> DNA
<213> Homo sapiens

<400> 235
tacagagaga gcctcggctt tgcctcatct ctaaagttgg aattgtagca gtgatactta 60
cctcataaag ctgttgagaa gactaaatta tttaatcaac atataaccct tatcaccata 120
tatgataggt tttataatat tccccagtga gataatggcg acaagaagta agggggcatt 180
tattaattgc tatattatat tactgctgac gtttcttatg atcagaactt tttacaatct 240
catggaatat tattgtccca ccttactaat aaggaaactg atgtctaata ccaagatact 300
ttgatttgca cat 313

<210> 236
<211> 567
<212> DNA
<213> Homo sapiens

<400> 236

ttctattata aaatgcttac gtgtgttcca gagcgactat tccagtgtca ccatttaatc 60
agaatgacat gcctgttcat gattctagaa ttcaggttgt ttaaatatga cagtaaccta 120
tgcagtcatg tgatcattaa tcacccacaa gttcagggac gccaaagatg atccagaggt 180
aggcatctgg ctatttcaac ttcacgaggc caaaatggaa gttaaagagc taatctgggc 240
tgaacagaag attctggcag aacactgaca ctgaatccaa ccttcacgca tcctctctcc 300
acatcattct ttaggcctta actctttatt ccacattctc taataaacct ctgtagctct 360
tcccacccac agagacagca tcctgagagc tctggggaac actcagatgt cctaaccaca 420
tgttaacctt gaattcccca gttccagtgg ttttcacctc tgctccacat gagataaccc 480
tttcaaagag ccatgcccag gatcttgtca tcagctgcac cacgctgcct agaaactgtg 540
agcttggacc attcactttg tggcttt 567

<210> 237
<211> 623
<212> DNA
<213> Homo sapiens

<400> 237
tttcgtttta agactgagta atcaggtttc tgttaatatt aagaggttaa aaatgattct 60
agtattcaat atgagcttcc agcctgaagt tttatttaat ttcaggttat aaccttaaat 120
aaacaatttt cctagcctca ttaatgctta atctaatgtt gcataacaaa gtatggactt 180
ccttagtcgg ttgatgctac tgagaatgtg caaatgtgtc acggccacat accagtacat 240
tagaagatct ctgtttttaa accttgtccc tcttctacaa acgttgtcta tcctcactgc 300
tcacagtgtg ctcctcaggc cagcactctc atcactggtt aaaatggaag attctcaggc 360
cctgtccttg tctcttgagc cagaatctgc attttaacgg acccccccaga ggattccaaa 420
gtactgtaaa gtttagaaag caccgctatt tacagtaatg cagacttgca gaattctgag 480
cctgcatagt atatcattct taaaatgata ctgtctttat tctgttaaaa ttgtaaagtt 540
tcactattag ggcttgattt tgtttttcta attattgcaa aatatttcaa acagaaaagt 600
atcagggctt atttccattg tca 623

<210> 238
<211> 506
<212> DNA
<213> Homo sapiens

<400> 238
aattcccggg tcgacgattt cgtccgactt ggcctgaagg aagtgacatg agcataagga 60
agtggggtgc ggtgggagcg ggtcaagtga cagaggaggg ctgggaaata ttaatgaata 120
gagtctgttg agaatcagcc aatgtgactt gaagcaacag gcaaccagaa tcgtgatggg 180
tcagtatttc ccccaggcct atccagacca ggtcttatcc tcctagcaac gcggccctct 240
cggctcttaa ggcgtgactt gccatggatg ttctgccttt ggaaccagtg ggtggtcaca 300
gcacagaggc tgttggtttc atggctgtca catgcacaga gacagccctg tccactgagc 360
ctgttctgtg gcagaaggaa ccccctagcc tggaccatct ttgggtggaa acaccagccc 420
ctcacgtctg actgtcattt tcagatgtga agagaacaca cagtgttgcc gatgtgtttg 480
tattgcctgt aaacaacgat tgtatt 506

<210> 239
<211> 385
<212> DNA
<213> Homo sapiens

<400> 239
ggcacgagga tcttttgttg cccttccttt ttaagccttt cctgtgcata gtgaagaata 60
ccatgactac tagtagtttg gtgttaccgc ctttatttgt gctaaagtgc cagcgttttt 120
acccaccatt gtatctgcac ccttacagca tatgtcaaca tgttagtatt cttgtaaaaa 180
tagtttggac ctgggggtct gaggtcccca ctttgggaac cattgaaata ggtacttaga 240

tctacagtgt atcatatctt ttcatctaca agattaaaaa catgatttca gttaattgtt 300
tttgtaattt tcttaatatg gttttgaggg gtttcagtcc agagtgaaga catgtatgtt 360
tcttggctta tgctgaagtt tactt 385

<210> 240
<211> 389
<212> DNA
<213> Homo sapiens

<400> 240
tttcgtccct ctgtaaggac gttttgtggt ggagccaggc aggagatgca aaggttggtc 60
cccaggatgt gcaggaagga ggctgagtgg ggagggccag gaagccgaag gcctgagcag 120
gaccccgggc atcacctctg atgcttgcat ggaggctttt gtgcccatgg ggtcctgggc 180
tccctacaac cacagcacga tctggggaga gaacggagag aagggagagg gtgcgcacag 240
ccagcccccg taaaatcctc ttcaaaacgc agccacccag gggttctagc acagacagat 300
gcccatgggg caggcagtgc ctgcatggga caggcacctg ccatatgccc aacagataac 360
aggacaactg tttacccatc cctggtatt 389

<210> 241
<211> 804
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(804)
<223> n = a,t,c or g

<400> 241
gaaaaataag atgtctacat atgtgctttc cttctactac tgaagactat ttaaaaacta 60
accttcactt aaaaaatact gaacatacaa ctaagtaaca ttaatgtcta ggtctaggtt 120
actaagttct cagtaaaaat gatgtttagt aaggaaatga tgtcagtcag cctgcagctc 180
atattgataa gccatccaaa attggtccca taggtgatga catctagacc tcacttcttc 240
cgctacctgt gctctcttcc ccctcttctc tttcctctgt tgncacaaag ccagctactg 300
ccaggatccc ccctccccat cgctctgcag tccagggtgg ggtctctgtt agcatgagga 360
aacgtgaaag gaaatgtatc tctgtgtacc tttatttttcc ctccagggct taaatggaaa 420
aagagagaag gtggagtccc ttttgccatt gtaactggaa ttaaaaacct taaaaaactc 480
aagttggagc tctgagtgca ttttcccaca aaaaccatgt tatatatagt ggtttggttc 540
tatggtacta ttaataagac agaatacttc agctatgatt tctgggctgg cctgggaaac 600
cactaaatat gacatagttt cagtgaaaaa gtgcgttcag ggttcccagc aatgagtttt 660
acaaacccac aatcagaaca caacccatct gaaagctgag ggccgcctac ttcgcagtag 720
gtttgacagc ctcacaaagg catccgaatg tcatagacgt caacgccaag cttaagagtc 780
tcctgcgtgg cgctccagca gtac 804

<210> 242
<211> 386
<212> DNA
<213> Homo sapiens

<400> 242
tttccctgtt ctgatgcccc agccccgcgt cacagcccgt tctcagtctc ctctgccttc 60
tctagtagat attgttagcc aggattccat tatcctcaat gctcccttcc ttcctgccac 120
agtccttggg aaatctcatc cacacacttg gatttttgct gataattcac aaatatatgt 180
cagctttcaa aaaccgcaca gatgagttta tgaatatggg tatgcagccg tatataaaat 240
ctccttaccg actatcaatg tctcagatca gcctgaaatt tgacttgtct caaactgatc 300

ttatttacc ccacaaattc tactctcctt cctccttccc tactgtcatg ttattctatt 360
cattcggtag gttaagccat aagcct 386

<210> 243
<211> 279
<212> DNA
<213> Homo sapiens

<400> 243
gataaaacag gggaaggagg gaacaggtag ggctggctgc aggaaagctt gcaggtatgg 60
ggtacacagt taaaagggag tgtcttctga tagtcttgat gcaggcctgg aggtcgtttg 120
ttatgggggt ggaggtcctg atgtacattg tggcggtaag atgccgtgct gtctttgcca 180
ccagcctgtg gcagccctgg tgttacacca gggcaggtgg gcagtttaat gtatctcagg 240
ccagatgacc tgccctgccc acctggtcga cgcggccgc 279

<210> 244
<211> 340
<212> DNA
<213> Homo sapiens

<400> 244
attttttttc atggttgtcc tatatattgt cagagcatat aatcattaca tactatgctg 60
tctctcctcc agtctttact tagtcttaat tctgttagta accgtatatt taatgctcac 120
cactagttct tataatgatg tttctttagt catttggatt gcctcaagtt ttgcctctag 180
taaattcttt aggaaaggct taagggaata ttcatatttc atgaattttc ttgctcgttc 240
ataacatttt gtctgtgttt tttatatcta aagtttgtgt ttgccagaaa tagaatcctt 300
ggctcacttt ttctttttctt gaatatctta aatttggaac 340

<210> 245
<211> 380
<212> DNA
<213> Homo sapiens

<400> 245
tccaccttca tgttaccctt ggtctcccca gtctctgcca gatccttctg acctggcatc 60
accaagacaa atgctggagt ggcctctcct tggacaaatc cttcccatga tcatcccact 120
acctcctctc cctgccttag tggtttggcc cattgggctc actcactgcc cttggccctc 180
ccctttttgtt ccagcctctc tggatgggtt ctacaacagt cgcagcctgg atggtaagtc 240
tccacccttg aggccagaga agtggtctcc ctggtcctgg ttctcacggc tcccaggtca 300
tgggcttccc aaggagggag ggcagagaaa gagagaggtt tgggcctagc aggttaaaaa 360
gctccttcgt cgacgcggcg 380

<210> 246
<211> 396
<212> DNA
<213> Homo sapiens

<400> 246
ctttgaaagg gccacatgct ctgccatgct agctgaactg cataagaatg catcagttat 60
gcaaagcatc ttttaccaga attgcacaaa tactagtttt tatttaaacc cttggctcat 120
gattaatgta tccatctaac aaagtcccaa aataatggaa ggatgatgct gtggcaggtt 180
taccctggac catctgcggc ggtgctctgc ctctttctcc atccaccctg gtccaggtcc 240

acagcagtgg agagagaaaa gagacagaag gatggccgtg gtcagcgcat gttgctgccg 300
cagccccagt gcctgatgtc tagttgttgc ttggtagacg tccagagttt aacaggatag 360
attcttgata ttctgcttga atcaacaatc ttacct 396

<210> 247
<211> 363
<212> DNA
<213> Homo sapiens

<400> 247
gttaatttct gattctgtga aggccagatc atacagctgg aaatgattga aaagaaaagc 60
atgttctttt gtttaacaag acaaagccta ttgtgtactc tgcttctgat gcttaaaaga 120
tgtatttttt tttcttactg cgtaatatgt agggctaaaa gttttgaact cttcacttca 180
gaaattactt ttccagacaa aagagcaaaa caatgtttgt ttaaattgtt ctctggaaca 240
tgactgatag tgtttacaac atttctatga cattttgatc ctgtggggag tacttggaga 300
ggtatttaaa gtgcttttca ctaaatcaga tgctttgtag tgtttgaacc ttttttaaaa 360
aaa 363

<210> 248
<211> 1120
<212> DNA
<213> Homo sapiens

<400> 248
tttcgttgaa ttattggcca cttggctggt aattacaaat ggtgtgggtg gtaggcagtg 60
gctgtcagtg ttcagtgctt taatgccatt gtgtaatgga tcagtaatga accttaaggc 120
tctttctggg tttgtgtgat aaaccttttt atggctacac agattttctt actgatgtag 180
caaagattgt aactgttctc attttgaaag caaagatgcc catacaacat gggaaatact 240
gatattttgt tgctgttgtc gttgttttgc ttcagttatg agttggttgc tggtaaaact 300
aaagcccagt ttggggttcc ctttgcagaa ttttctgttt tcttgatttt ggaaaatgtt 360
actgcttgtc gttttctgta tatttaactg gagctgggga ataattggtg tttggcaatc 420
tctgtatcat tcattgtagg caaagctgct actgtcaccc acagatacat ttgtttgcat 480
tccattgcaa ggattgaagt cagcacattc ttagtcccaa attgcttttg gatattaatg 540
tttatcactt ctataaccat aaggagaatc aaaaggctat tggaaaaagt caaagcttgt 600
tatgcgcctt acatattggg tccgagaacc cattaaaatg cccgaggtct aaaatacccc 660
ccagaagttg cactatcatt tcttccttct ttcttccccc ggcaacgcta tttcatttgt 720
cccctcccgc gtcctgccgc agttcacaaa cactccttcc ctgcgcctcc agcgccgcca 780
tgcacccata ctctccatcc cttctcgtgc cgggtcgtgg caacacacat ctcgcgtccc 840
ccccccccgcg ttccgtcccc ccccacacac tcccctctcc ccctcactcc cggtatccag 900
ccctccccac actacaccgc tggggcccgg cgccccgccc ccccccctcc cccgtctcgg 960
ccgcggcccc acctggacac cacacaccct ctcgcccgcc ccccccctcc tcactttgcg 1020
caaaacccca catacccccg ctccctcgcc cccggatcac agcccgatca ccaactcacc 1080
ccgcgcgccc ctccgcgtcg ccccgcgcgc cacaccgccc 1120

<210> 249
<211> 353
<212> DNA
<213> Homo sapiens

<400> 249
cgcggccgcg tcgacctagg cagagaaacc atgagctgga taatctttcc ttgtctttga 60
agtctctgtt actgagaact ggtgagaaca ggtgagcaaa gacaaatgaa tctctgccta 120
agagcctatt gctgttgggg actacatgtc catattggtt gtttctgcct ttctggctaa 180
tttgtggctg ttgatgacca tctcgacttc ccagatgtta aatatgacaa aaataaccta 240
tcttgtcctt tttcttcacc tttcagctct cagaattggc agtacacccc acagcttctt 300

gctaaagtcc tatcacctgg ggacccactt ttcattattt cacatgaact cac 353

<210> 250
<211> 372
<212> DNA
<213> Homo sapiens

<400> 250
cacaaggttt tatcatatca tagagatgcc attttagaga tggttataga cattaagcaa 60
gtaaattaat ttcacaatac aagcaagatg tgacaaatgc tatgaaagaa ataagacaga 120
atgatataat ggagtgactg ggaaaaggag gacatgtgaa tttggagttc aggaaagtca 180
tttataagca ggagatattt gaggctatac aaaggtctgg gtacagagca atctggacag 240
cagtacagca agcacaaagg tcttaaaatg ggaacaagct cggtatggat ggagacaaat 300
acaaacattg gaaggaataa agttaccaac acagctttcc cacaggtgat gagaaagccc 360
aactgagaga aa 372

<210> 251
<211> 441
<212> DNA
<213> Homo sapiens

<400> 251
gccaaaattc ctaagaaaat attagcaaac cccctccatc gatatataaa aataatccaa 60
catgataaaa ttgggtttat tccagtaatt gcaaggttgg ctgaccattc agaaaactaa 120
tgtactataa ttgccacatt aaaaagtaaa gtaataatgt cttaaaggat tttaaaagta 180
aaggaaaaat atcacatagt gcctcaatag atgatgaaat ctgaagtcaa tttataatta 240
agacttagaa atgaacagtt tgataaaaga aaaactctaa agcaaacatc acacttagtg 300
gtggaatgct gaaatgttcc ctttgagatt cgaaaccaga ttaggatgat cactctcatc 360
atgtctattt aacattttgc tgtaaatcca agatgctgtg gtatgataag aaaaagacga 420
aatcgtcgac ccgggaattc c 441

<210> 252
<211> 342
<212> DNA
<213> Homo sapiens

<400> 252
tttccctgta tatagaaaac ggatttaaag acaactgttt aagtacctct ttaatatcaa 60
aacactaaat ttacaaattg cacatttatg caaatggttg taccaaggct tctctctgta 120
ccccaattgc tcaacacagc tccactgttc ttaccatggg agaaaaccgt aaaaactcag 180
tattctggaa taatctttaa attcaaaagt agaatagaaa cagcagagaa aagtataggg 240
gataccaaag agagaattca gccaagtcaa atatagaata cactgtcttc ctgcttttcc 300
tacctaaacc acattattgc cctattctta atcccccaa aa 342

<210> 253
<211> 349
<212> DNA
<213> Homo sapiens

<400> 253
ctttgtccat gctctgtcct cccctcccca ctgtaaatac tcttcttttg ggctaggggg 60
ctgtgccctc cagggtattg ttttccttcc tgttcagtgc ccatgctctt gcctgtgttt 120

ctcctgtacc tttcccagga cctagcagac agccgagcac ctgcccactg ctctgtaaat 180
actgacctgc atctcaagtg gggtagcctt tgtgtacttt cccatttcca agttgatttg 240
cctgtgaatc ccatatgtga acacatatgc agatgtccct gatgtgctct ttcctgcttg 300
gaggcacatg caaggatgtg gcacttttc taattgcttg aacagtcaa . 349

<210> 254
<211> 366
<212> DNA
<213> Homo sapiens

<400> 254
atttgagtca gttcagccct ttgagaaacc taggatcaag agttaataga ccatgaatct 60
agaacatgtc attgtatctc tcttaacatt ttatagagtc ctcctctata aagagataat 120
tggactgcat cattgttttc aacattttca tgttaatgca tttttactat caccctccc 180
cccatcatga ctattgccct actttgagta gacatagatt gagaatatgt tttggctaaa 240
accatcacat tcatcagtta cgagccactg tttactattt ctagaggtaa ttttgagaaa 300
attgtgaatg acagcacttc tctcctggga ttactaatta acacttgagc agatttgtac 360
aggcta 366

<210> 255
<211> 400
<212> DNA
<213> Homo sapiens

<400> 255
tgctcgtgtg gatatgatca cagctgagta catgggcttg tgcgtgcaca agaatagaaa 60
aggatgtgca tgtttgtgca tgcacgagga cacacagagc acacgagcac ccacaggaca 120
catagacaag ccaaagggca ctcatgtgtg ggggtcaaac atctgagagc gtgtgaccag 180
tgtgagtctg aagaccatag actctagtga caggtgtctg agatggccac gggctgcatt 240
agaaacacta agaggctcac tgcctgctct gtccccattg gatccctcgg caggctccag 300
tgacagcaaa ctgtggcatg gagcccccag aggacctagc acttaacctt ggagcttccc 360
tgagtcctca ccttgaagag agtccctgga ccaatgtcca 400

<210> 256
<211> 338
<212> DNA
<213> Homo sapiens

<400> 256
tttttttaa agaaaatccc agtgtgggta aaaatctaac aatgaaatat taaattgatt 60
acctgcatag cactgttcta attaataaac agtttctata aacatcaagg aacatcacct 120
agagcattaa tcaggtacca aacaatttaa ccattgactt cacacatatg catatgtagg 180
tggatgctca tacagtttac cttcttttt cccatttccc agagattttg gttttgtttg 240
ttgttttttt tcccccaaac ctttaagtgt atgaaattct attctttgat tgaagtcagg 300
gaaggcgtat gcattcacca gagaagtcaa tttgttca 338

<210> 257
<211> 336
<212> DNA
<213> Homo sapiens

<400> 257

agggcctaat gtgttttcag tatcggatga ctttcaagag tataataagc cagggaaaaa 60
atggaaggtt atgactcaga tttttattaa tcagctcaca gtggttcaca aaggacccca 120
gttctgtaaa ttcttcctgg aagggaggtg tattaaggga gatcagtgta attttgttca 180
tgatgcagag ttgtagaaaa gaaaagacat gtgcaaattt tattttcaag gatgttgtgc 240
caaaggaaag aagtgcattt atatgcataa tgaatttcca tgcaagttct atcatagtgc 300
agcaaaatgt taccagggag gcttctgtga attttc 336

<210> 258
<211> 358
<212> DNA
<213> Homo sapiens

<400> 258
atcactaaga agatgaatta acttagtctc tgatatttat ctaagatgga ctcctgtcaa 60
aaatctctaa ccagtcctcc atgcccaacc gcatagaact ctctgctctg ttcagggtac 120
catttcatgt tgattggctt ttatttgctt ttgtttttt ccctctgtgt tatgtccttg 180
ggttctgcct accttccaca gtcaaaacac attcgttgat tgtgggtgct gtttttcaaa 240
gtgagattgt cactaatgtt tgccagattt ataaatgtaa aacaacttga cagattttgt 300
atgagttttt tgggtgcaga tgacttaagt tctggaatta ttaacatgtg actgagag 358

<210> 259
<211> 1092
<212> DNA
<213> Homo sapiens

<400> 259
tattgcttgt tcgacgattt cgctgcacag actctttgct gcatggctca ctacctgtgg 60
tgttggaggg gtggttttga tgcttcagga ctgtctttcc tgttctcctc aatgtctctt 120
tcaagatatg aagttaaaat taggtactgg gattgctcac ctgagttttg gtccttgaga 180
tgatgctttt ctgtgtgcag atagttgcta aaatttggtg ttcctgctgg gggtacaaac 240
tatgtaggtt tctattctgc cattttgttc tatcctttat caactattta ttttaggaaa 300
aggaatgtgt gtacaattgg agagccattt acaaaatctt tactacaaat gtttttgatt 360
ttataatttg tgataacagg aactatactt ttataactct ataaacctct gggtgtccct 420
aatacataga atagtgagag ctcgagagtt agttttagtt gcctttccag gtaggaccat 480
cagaaggaat gacagagcag tttgttttga aggtagacta aagcaccaag gtctaggtgt 540
cacggaaatt acgttttact ggatcgagtc ttgctctgct cggtctttcc cttttaccct 600
acatttcctc ctctcgatct tccgtgcagc gtaccacgcc catctgactc gagagctccc 660
cacctgcaaa cctttatctt ccttgttgcg gctgcctagc ggatcccctc gccaccgagg 720
ccagattctg cttgttcaga acttctgccc gtcttccgtc ttcttttaac gacgcacacc 780
cggccacgcg cggacggacc tatcccctcg ccccggtcca ccggggcggc gctggactta 840
caccctttct ccgcgctcct ggggccatgc gcgcgcatcc ccatacccca cgctccgccg 900
cgcacattat actctctact tggcttgcgc tcgggcccta tctctatcgc actcgtaact 960
gcggtccctt gactgaagaa ccatccctcg caccttctac tcagggcgcg cggcgtacta 1020
acactactcc cgagctcgcg gcccgtccgg tcgctccgtg ccttcatata tcttacacgt 1080
tcacgcgcgg cc 1092

<210> 260
<211> 317
<212> DNA
<213> Homo sapiens

<400> 260
cttcctctga ctcttgctgg tttacttcct gctataatcc tctaaatatg gacttgaccc 60
agagctcagc acacaattaa ctgctcttgc tttacattat atgcctttgc tttacattat 120
atgccttaga cagcttgttc tgttccattc caaatgccac tctcaacact cctgcagagc 180

gggaggcata cagtactcaa tgcacgtgag tctctttctt tcttcaccca taaactatga 240
caatggtttt ttagtgtctc ccaccttccc tttacatgtc aaactgagct tcctaaaata 300
ctcctttaaa tgcatca 317

<210> 261
<211> 1187
<212> DNA
<213> Homo sapiens

<400> 261
aattcctggg tcgacgattt cgtaagatcg atagtatctc aggcaagtaa agtggcaatg 60
agatacagca gagtagacag caacttttac atatgtattt tggaggtgga atcagaagga 120
tttgcttaat gggttggctg tggagggaga gggaggaatc aagggtgcct tcaggttgca 180
ggctccagcc ttgggctgct gatggtacca tttgctggtt ggagaagact ggaggaggaa 240
caggtggaga ggagtgggta gcaagagttc tgtgtaaata cttgggaggc atccaagcgg 300
agagttaagt aggcactgaa tatttaagtt gagctgaggg gagtgatcta gactggacat 360
aaattttggg agtcactagt atacagatgg catgtcatgg aactgattga gattgtttgt 420
ggccttaaga tcaagccctg cgagactgga gtaataaaac tttggtctcc cacacagcca 480
gctctgtgtg gggaaaaaaa aaagccctaa aacactaaca acggttaaag cttgggcaaa 540
ggagactgaa aaggttcagc cattaaaggg ggagagtatt ttattatttt caagaaagag 600
ggaatggtca cctctgtcaa atgctgctga gaagttacac aatgagaata gagaaatgtc 660
tatttggatt tgacaacatg atggtgactg ttttgacaag tgcgccaagc cacattggga 720
tgcttcgaag agaaaatagg aaatggggtg aatatcgaca gcttcgttgg gaaattttgc 780
tgctggaaag tggaagaaaa acacttatgg gcttcgaggg agatggggcc aaaaagaggg 840
ctttttttt ttatttttt tttagacaca gagccttcag ttctcggcgg gggggcgggg 900
gggcaatacg gtttaaccag aactcacctt tggcgccagg gggcctccta cttgacgctt 960
ctgagaaacg gaacacagag cggccccctg gcccacttat aaaaaaaaat tcttttatca 1020
caggttctcc tgtacggaga gcttttgtta acagcggact caagcccttc gagtagtacc 1080
gccgggcctc gccgcaaaaa caaaacagtc tccttctttt ctttcttgac aggggccgtc 1140
tgcccccgc ccatctcatc ctatacgtac gccggcgtgg gcggccg 1187

<210> 262
<211> 69
<212> PRT
<213> Homo sapiens

<400> 262
Met Cys Ile Glu Leu Cys Ile Ile Phe Pro Val Val Met Ser Ser Pro
1 5 10 15
Ser Phe Phe Ile Leu Val Ser Glu Asn Asn Asn Asn Asp Leu Phe Leu
20 25 30
Ile Leu Met Val Leu Leu Leu Val Ser Leu Gly Leu Thr Leu Ala Val
35 40 45
Leu Val Asp Trp Leu Gln Leu Glu Asp Leu Lys Met Ala Ser Pro Ile
50 55 60
Cys Leu Gly Pro *
65 68

<210> 263
<211> 97
<212> PRT
<213> Homo sapiens

<400> 263
Met Met Gln Leu Leu Phe Pro Leu Pro Leu Trp Gly Ile Ile Pro Phe

```
 1               5                  10                  15
His Leu His Cys Cys Asp Ile Val Cys Pro Leu Ser Gln Val Glu Gly
            20                  25                  30
Gly Ile Leu Arg Leu Pro Pro Ala Leu Val His Ser Ile Phe Leu Leu
        35                  40                  45
His Ala Ala Cys Val Met Ile Ser Cys Lys Ala Phe Asn Ile Lys Ser
    50                  55                  60
Pro Leu Cys Val Arg Pro His Val Glu Arg Lys Thr Cys Leu Arg Glu
65                  70                  75                  80
Glu Val Cys Cys Val Ser Pro Phe Ser Ser Pro Gln Ile Cys Val Ser
                85                  90                  95  96
 *
```

<210> 264
<211> 50
<212> PRT
<213> Homo sapiens

<400> 264

```
Met Ser Arg Phe Phe Ile Phe Cys Cys Leu Arg His Phe Ser Tyr Phe
 1               5                  10                  15
Ser Asp Phe Ala Val Leu Phe Leu Gly Ala Leu Glu His Leu Lys Tyr
            20                  25                  30
Gln Leu Ala Val Gly His Ser Val Leu Ser Glu Ser Thr Asp Gly Asn
        35                  40                  45
Val  *
 49
```

<210> 265
<211> 51
<212> PRT
<213> Homo sapiens

<400> 265

```
Met Ala Tyr Arg Gly Gln Leu Leu Ala Gly Phe Thr Phe Asp Val Ser
 1               5                  10                  15
Ala Cys Leu Trp Thr Ser Trp Arg Thr Ala Leu Thr Glu Cys Val Ala
            20                  25                  30
Trp Gly Ile Cys Pro Leu Gly Trp Val Val Pro Val Leu Gly Pro Val
        35                  40                  45
Asp Gly  *
    50
```

<210> 266
<211> 64
<212> PRT
<213> Homo sapiens

<400> 266

```
Met His Arg Gly Val Leu Val Thr Leu Leu Lys Ile Thr Val Leu Lys
 1               5                  10                  15
Ser Met His Arg Gly Ile Leu Val Thr Leu Leu Lys Ile Thr Ile Leu
            20                  25                  30
Lys Ser Met His Arg Gly Val Leu Asp Thr Leu Leu Lys Ile Thr Ile
```

```
         35                  40                  45
Leu Lys Ser Met His Arg Gly Val Leu Val Thr Leu Leu Lys Ile Thr
     50                  55                  60                  64
```

<210> 267
<211> 202
<212> PRT
<213> Homo sapiens

<400> 267

```
Met Val Met Leu Ala Ala Leu Ala His His Leu Phe Tyr Trp Asp Val
  1               5                  10                  15
Trp Phe Ile Tyr Asn Val Cys Leu Ala Lys Val Lys Gly Tyr Arg Ser
             20                  25                  30
Leu Ser Thr Ser Gln Thr Phe Tyr Asp Ala Tyr Ile Ser Tyr Asp Thr
         35                  40                  45
Lys Asp Ala Ser Val Thr Asp Trp Val Ile Asn Glu Leu Arg Tyr His
     50                  55                  60
Leu Glu Glu Ser Arg Asp Lys Asn Val Leu Leu Cys Leu Glu Glu Arg
 65                  70                  75                  80
Asp Trp Asp Pro Gly Leu Ala Ile Ile Asp Asn Leu Met Gln Ser Ile
                 85                  90                  95
Asn Gln Ser Lys Lys Thr Val Phe Val Leu Thr Lys Lys Tyr Ala Lys
            100                 105                 110
Ser Trp Asn Phe Lys Thr Ala Phe Tyr Leu Ala Leu Gln Arg Leu Met
        115                 120                 125
Asp Glu Asn Met Asp Val Ile Ile Phe Ile Leu Leu Glu Pro Val Leu
    130                 135                 140
Gln His Ser Gln Tyr Leu Arg Leu Arg Gln Arg Ile Cys Lys Ser Ser
145                 150                 155                 160
Ile Leu Gln Trp Pro Asp Asn Pro Lys Ala Glu Gly Leu Phe Trp Gln
                165                 170                 175
Thr Leu Arg Asn Val Val Leu Thr Glu Asn Asp Ser Arg Tyr Asn Asn
            180                 185                 190
Met Tyr Val Asp Ser Ile Lys Gln Tyr  *
        195                 200 201
```

<210> 268
<211> 87
<212> PRT
<213> Homo sapiens

<400> 268

```
Met Leu Phe Trp Leu Ile Lys Val Ser Cys Ser Phe Ser Cys Ser Asp
  1               5                  10                  15
Glu Thr Ser Ala Ala Ser Trp Gly Phe Gly Ala Phe Ser Phe Ser Phe
             20                  25                  30
Leu Leu Leu Gly Ile Ser Cys Leu Met Arg Leu Val Pro Asp Thr Phe
         35                  40                  45
Val Leu Phe Ser Phe Ser Cys Glu Leu Phe Ser Cys Phe Arg Gly Leu
     50                  55                  60
Ile Gly Gly Arg Gly Leu Ser Ser Ser Pro Leu Ile Asn Leu Ser Tyr
 65                  70                  75                  80
Gly Arg Ile Asn Leu Ser  *
                 85  86
```

<210> 269
<211> 343
<212> PRT
<213> Homo sapiens

<400> 269
Met Leu Gln Gly His Ser Ser Val Phe Gln Ala Leu Leu Gly Thr Phe
1 5 10 15
Phe Thr Trp Gly Met Thr Ala Ala Gly Ala Ala Leu Val Phe Val Phe
20 25 30
Ser Ser Gly Gln Arg Arg Ile Leu Asp Gly Ser Leu Gly Phe Ala Ala
35 40 45
Gly Val Met Leu Ala Ala Ser Tyr Trp Ser Leu Leu Ala Pro Ala Val
50 55 60
Glu Met Ala Thr Ser Ser Gly Gly Phe Gly Ala Phe Ala Phe Phe Pro
65 70 75 80
Val Ala Val Gly Phe Thr Leu Gly Ala Ala Phe Val Tyr Leu Ala Asp
85 90 95
Leu Leu Met Pro His Leu Gly Ala Ala Glu Asp Pro Gln Thr Ala Leu
100 105 110
Ala Leu Asn Phe Gly Ser Thr Leu Met Lys Lys Lys Ser Asp Pro Glu
115 120 125
Gly Pro Ala Leu Leu Phe Pro Glu Ser Glu Leu Ser Ile Arg Ile Gly
130 135 140
Arg Ala Gly Leu Leu Ser Asp Lys Ser Glu Asn Gly Glu Ala Tyr Gln
145 150 155 160
Arg Lys Lys Ala Ala Ala Thr Gly Leu Pro Glu Gly Pro Ala Val Pro
165 170 175
Val Pro Ser Arg Gly Asn Leu Ala Gln Pro Gly Gly Ser Ser Trp Arg
180 185 190
Arg Ile Ala Leu Leu Ile Leu Ala Ile Thr Ile His Asn Val Pro Glu
195 200 205
Gly Leu Ala Val Gly Val Gly Phe Gly Ala Ile Glu Lys Thr Ala Ser
210 215 220
Ala Thr Phe Glu Ser Ala Arg Asn Leu Ala Ile Gly Ile Gly Ile Gln
225 230 235 240
Asn Phe Pro Glu Gly Leu Ala Val Ser Leu Pro Leu Arg Gly Ala Gly
245 250 255
Phe Ser Thr Trp Arg Ala Phe Trp Tyr Gly Gln Leu Ser Gly Met Val
260 265 270
Glu Pro Leu Ala Gly Val Phe Gly Ala Phe Ala Val Val Leu Ala Glu
275 280 285
Pro Ile Leu Pro Tyr Ala Leu Ala Phe Ala Ala Gly Ala Met Val Tyr
290 295 300
Val Val Met Asp Asp Ile Ile Pro Glu Ala Gln Ile Ser Gly Asn Gly
305 310 315 320
Lys Leu Ala Ser Trp Ala Ser Ile Leu Gly Phe Val Val Met Met Ser
325 330 335
Leu Asp Val Gly Leu Gly *
340 342

<210> 270
<211> 66
<212> PRT
<213> Homo sapiens

<400> 270
Met Lys Cys Lys Leu Ile Pro Val Cys Pro Phe Leu Arg Leu Asn Thr

```
  1               5                  10                  15
Gln Pro Leu Leu Ile Ile Ser Tyr Gly Ile Phe Leu His Ile Phe Arg
            20                  25                  30
Asp Phe Ser Tyr Ile His Arg Val Arg Glu Arg His Ser Val Phe Leu
        35                  40                  45
Ser Val Gly Gln Gln Trp Cys Pro Glu Leu Thr Arg Ser Ile Phe Leu
    50                  55                  60
Leu Asn
 65  66
```

<210> 271
<211> 209
<212> PRT
<213> Homo sapiens

<400> 271

```
Met Gly Leu Gly Ala Arg Gly Ala Trp Ala Ala Leu Leu Leu Gly Thr
  1               5                  10                  15
Leu Gln Val Leu Ala Leu Leu Gly Ala Ala His Glu Ser Ala Ala Met
            20                  25                  30
Ala Ala Ser Ala Asn Ile Glu Asn Ser Gly Leu Pro His Asn Ser Ser
        35                  40                  45
Ala Asn Ser Thr Glu Thr Leu Gln His Val Pro Ser Asp His Thr Asn
    50                  55                  60
Glu Thr Ser Asn Ser Thr Val Lys Pro Pro Thr Ser Val Ala Ser Asp
 65                  70                  75                  80
Ser Ser Asn Thr Thr Val Thr Thr Met Lys Pro Thr Ala Ala Ser Asn
                85                  90                  95
Thr Thr Thr Pro Gly Met Val Ser Thr Asn Met Thr Ser Thr Thr Leu
            100                 105                 110
Lys Ser Thr Pro Lys Thr Thr Ser Val Ser Gln Asn Thr Ser Gln Ile
        115                 120                 125
Ser Thr Ser Thr Met Thr Val Thr His Asn Ser Ser Val Thr Ser Ala
    130                 135                 140
Ala Ser Ser Val Thr Ile Thr Thr Thr Met His Ser Glu Ala Lys Lys
145                 150                 155                 160
Gly Ser Lys Phe Asp Thr Gly Ser Phe Val Gly Gly Ile Val Leu Thr
                165                 170                 175
Leu Gly Val Leu Ser Ile Leu Tyr Ile Gly Cys Lys Met Tyr Tyr Ser
            180                 185                 190
Arg Arg Gly Ile Arg Tyr Arg Thr Ile Asp Glu His Asp Ala Ile Ile
        195                 200                 205         208
 *
```

<210> 272
<211> 105
<212> PRT
<213> Homo sapiens

<221> misc_feature
<222> (1)...(105)
<223> Xaa = any amino acid or nothing

<400> 272

```
Met Trp Leu Pro Pro Ala Leu Leu Leu Leu Ser Leu Ser Gly Cys Phe
  1               5                  10                  15
Ser Ile Gln Gly Pro Glu Ser Val Arg Ala Pro Glu Gln Gly Ser Leu
```

```
            20                  25                  30
Thr Val Gln Cys His Tyr Lys Gln Gly Trp Glu Thr Tyr Ile Lys Trp
        35                  40                  45
Trp Cys Arg Gly Val Arg Trp Asp Thr Cys Lys Ile Leu Ile Glu Thr
    50                  55                  60
Arg Gly Ser Glu Gln Gly Glu Lys Ser Asp Arg Val Ser Ile Lys Asp
65                  70                  75                  80
Asn Gln Lys Asp Arg Thr Phe Thr Val Thr Met Glu Gly Leu Arg Arg
                85                  90                  95
Asp Asp Ala Xaa Val Tyr Trp Cys Gly
            100                 105
```

<210> 273
<211> 61
<212> PRT
<213> Homo sapiens

<400> 273

```
Met Lys Phe His Leu Ser Phe Phe Ser Leu Lys Arg Ala Ile Phe Tyr
 1               5                  10                  15
Ile Cys Ala Lys Ala Asp Lys Ile Ser Gly Gly Tyr Leu Tyr Lys Cys
            20                  25                  30
Arg Thr Val Ser Tyr Ser Gly Lys Asn Val Arg Ser Gly Val Lys Ile
        35                  40                  45
Ser Gly Phe Leu Ser Ala Cys Ile Ile Ser Tyr Leu  *
    50                  55                  60
```

<210> 274
<211> 149
<212> PRT
<213> Homo sapiens

<400> 274

```
Met His Met Leu Asn Gly Ala Leu Leu Ala Leu Leu Phe Pro Val Val
 1               5                  10                  15
Asn Thr Arg Leu Leu Pro Phe Glu Leu Glu Ile Tyr Tyr Ile Gln His
            20                  25                  30
Val Met Leu Tyr Val Val Pro Ile Tyr Leu Leu Trp Lys Gly Gly Ala
        35                  40                  45
Tyr Thr Pro Glu Pro Leu Ser Ser Phe Arg Trp Ala Leu Leu Ser Thr
    50                  55                  60
Gly Leu Met Phe Phe Tyr His Phe Ser Val Leu Gln Ile Leu Gly Leu
65                  70                  75                  80
Val Thr Glu Val Asn Leu Asn Asn Met Leu Cys Pro Ala Ile Ser Asp
                85                  90                  95
Pro Phe Tyr Gly Pro Trp Tyr Arg Ile Trp Ala Ser Gly His Gln Thr
            100                 105                 110
Leu Met Thr Met Thr His Gly Lys Leu Val Ile Leu Phe Ser Tyr Met
        115                 120                 125
Ala Gly Pro Leu Cys Lys Tyr Leu Leu Asp Leu Leu Arg Leu Pro Ala
    130                 135                 140
Lys Lys Ile Asp  *
145             148
```

<210> 275
<211> 258

<212> PRT
<213> Homo sapiens

<400> 275

```
Met Arg Trp Ile Ala Phe Ala Val Met Ile Val Leu Ala Leu Ile Arg
  1               5                  10                  15
Ile Gly His Gly Gln Gly Glu Gly His Pro Pro Leu Ala Asp Phe Ser
            20                  25                  30
Gly Val Arg Asn Leu Phe Gly Val Cys Val Tyr Ser Phe Met Cys Gln
        35                  40                  45
His Ser Leu Pro Ser Leu Ile Thr Pro Val Ser Ser Lys Arg His Leu
    50                  55                  60
Thr Arg Leu Val Phe Leu Asp Tyr Val Leu Ile Leu Ala Phe Tyr Gly
65                  70                  75                  80
Leu Leu Ser Phe Thr Ala Ile Phe Cys Phe Arg Gly Asp Ser Leu Met
                85                  90                  95
Asp Met Tyr Thr Leu Asn Phe Ala Arg Cys Asp Val Val Gly Leu Ala
            100                 105                 110
Ala Ala Arg Leu Phe Leu Gly Leu Phe Pro Val Phe Thr Ile Ser Thr
        115                 120                 125
Asn Phe Pro Ile Ile Ala Val Thr Leu Arg Asn Asn Trp Lys Thr Leu
    130                 135                 140
Phe His Arg Glu Gly Gly Thr Tyr Pro Trp Val Val Asp Arg Val Val
145                 150                 155                 160
Phe Pro Thr Ile Thr Leu Val Pro Pro Val Leu Val Ala Phe Cys Thr
                165                 170                 175
His Asp Leu Glu Ser Leu Val Gly Ile Thr Gly Ala Tyr Ala Gly Thr
            180                 185                 190
Gly Ile Gln Tyr Val Ile Pro Ala Phe Leu Val Tyr His Cys Arg Arg
        195                 200                 205
Asp Thr Gln Leu Ala Phe Gly Cys Gly Val Ser Asn Lys His Arg Ser
    210                 215                 220
Pro Phe Arg His Thr Phe Trp Val Gly Phe Val Leu Leu Trp Ala Phe
225                 230                 235                 240
Ser Cys Phe Ile Phe Val Thr Ala Asn Ile Ile Leu Ser Glu Thr Lys
                245                 250                 255
Leu  *
257
```

<210> 276
<211> 101
<212> PRT
<213> Homo sapiens

<400> 276

```
Met Ala Leu Ala Leu Ala Ala Tyr Val Cys Gly Trp Val Val Asp Arg
  1               5                  10                  15
Glu Thr Trp Pro Val Pro Met Pro Cys Asn Lys Gly Gly Arg Ala Cys
            20                  25                  30
Asn Leu Glu Met Gly Met Glu Trp Leu Asn Leu His Cys Glu Val Ser
        35                  40                  45
Lys Trp Gln Gln Pro Pro Ser Gly Ala Leu Cys Cys Ser Leu Ala Pro
    50                  55                  60
Leu Gln Ser Ile Phe Phe Pro Ala Ala Lys Val Met Phe Lys Asn Gly
65                  70                  75                  80
Ser Trp Thr Val Leu Leu Pro Cys Ser Glu Phe Pro Ile Gly Phe Pro
                85                  90                  95
Ser His Leu Glu  *
            100
```

<210> 277
<211> 82
<212> PRT
<213> Homo sapiens

<400> 277
Met Arg Cys Gly Trp Gly Pro Leu Gly Cys Leu Gly Thr Gly Ala Pro
1 5 10 15
Ala Gly Trp Met Val Leu Gly Ser Pro Arg Ser Gln Leu Gln Arg Ala
20 25 30
Arg Trp Ser Arg Ala Ser Leu Ser Ala Phe Gly Trp Glu Ile Arg Leu
35 40 45
Arg Pro Glu Gly Pro Lys Ala Pro Arg Gln Leu Leu Leu Val Ala Leu
50 55 60
Glu Ser Glu Thr Leu Gly Val His Gly Gly Ala Thr Pro Leu His Cys
65 70 75 80
Leu *
81

<210> 278
<211> 65
<212> PRT
<213> Homo sapiens

<400> 278
Met Asn Asn Ser Pro Leu Ala Leu Phe Ser Trp Glu Gly Trp Lys Lys
1 5 10 15
Phe Leu Val Leu Leu Pro Ala Phe Cys Ile Thr Pro Ser Gln Ser Thr
20 25 30
Ser Phe Ser Asn Ile Val Pro Thr Thr Tyr Gln Tyr Cys Thr Pro Gly
35 40 45
Ser Cys Gln Ala Val His Ser Asn Ala Val Gly Gly Asn Thr Trp Lys
50 55 60 64
*

<210> 279
<211> 89
<212> PRT
<213> Homo sapiens

<400> 279
Met Phe Ser Cys Phe Phe Ser Thr Ser Leu Ala Thr Ser Val Ser Leu
1 5 10 15
Glu Ala Gln Ser Cys Phe Ala Trp Pro Leu Ile Val Ser Phe Pro Gln
20 25 30
Gly Ser Leu Leu Ser Pro Phe Leu Leu Met Ser Tyr Asn Leu Ser His
35 40 45
Leu Ile Tyr Ser Gly Glu Leu Asn Gly Arg Leu Tyr Ala Glu Asn Ser
50 55 60
Gln Ile Cys Ile Cys Ser Pro Ala Phe Pro Thr Lys Leu Tyr Leu His
65 70 75 80
Ile Phe Ala Asp Leu Ile Thr Ser *
85 88

<210> 280
<211> 57
<212> PRT
<213> Homo sapiens

<400> 280
Met Cys Leu Ala His Leu Phe Lys Leu Leu Val Tyr Phe Asn Arg Ser
1 5 10 15
Asn Ser Trp Val Gln Ala Pro Phe Val Leu Glu Thr Thr Thr Gly Leu
20 25 30
Phe Ser Ser Ser Val Ser Leu Ile Cys Ile Leu Asn Leu Phe Cys Lys
35 40 45
Gln Asn Leu Asn Asn Asn Phe Leu *
50 55 56

<210> 281
<211> 65
<212> PRT
<213> Homo sapiens

<400> 281
Met Leu Pro Pro Leu Cys Trp Cys Cys Val Arg Thr Met Thr Cys Cys
1 5 10 15
Ile Gly Thr Ser Thr Gly Met Asp Gly Arg Pro Pro Ser Pro Trp Arg
20 25 30
Arg Ile Pro Cys Trp Thr Gln Thr Cys Ser Cys Arg Asn Ser Ala Thr
35 40 45
Pro Ser Ser Pro His Phe Leu His Thr Ser Arg Trp Pro Gly Pro Met
50 55 60
Tyr
65

<210> 282
<211> 78
<212> PRT
<213> Homo sapiens

<400> 282
Met Thr Ala Arg Phe Leu Ile Cys Leu Phe Gln Thr Thr Met Tyr Ala
1 5 10 15
Glu Phe Asn Leu Gly Gln Arg Arg Trp Gln Thr Arg Asn Ala Pro Asn
20 25 30
Leu Ser Gly Trp Leu Gly Leu Ala Gly Ala Ala Pro Trp Gln Gly Arg
35 40 45
Ile Ser Pro Met Leu Gly Thr Lys Val Ser Leu Cys Asn Leu Ser Glu
50 55 60
Glu Ser Leu Ala Pro Leu Ala Lys His Thr Pro Arg Ala *
65 70 75 77

<210> 283
<211> 61
<212> PRT
<213> Homo sapiens

<400> 283
Met Ser Pro Thr Gly Leu Leu Val Val Phe Ala Pro Val Val Leu Gly
1 5 10 15
Leu Lys Ala Ile Thr Leu Ala Ala Leu Leu Leu Ala Leu Ala Thr Ser
20 25 30
Arg Arg Ser Pro Gly Gln Glu Asp Val Lys Thr Thr Gly Pro Ala Gly
35 40 45
Ala Met Asn Thr Leu Ala Trp Ser Lys Gly Gln Glu *
50 55 60

<210> 284
<211> 61
<212> PRT
<213> Homo sapiens

<400> 284
Met Ser Pro Thr Gly Leu Leu Val Val Phe Ala Pro Val Val Leu Gly
1 5 10 15
Leu Lys Ala Ile Thr Leu Ala Ala Leu Leu Leu Ala Leu Ala Thr Ser
20 25 30
Arg Arg Ser Pro Gly Gln Glu Asp Val Lys Thr Thr Gly Pro Ala Gly
35 40 45
Ala Met Asn Thr Leu Ala Trp Ser Lys Gly Gln Glu *
50 55 60

<210> 285
<211> 66
<212> PRT
<213> Homo sapiens

<400> 285
Met Phe Phe Trp Arg Leu Leu Ile Leu Tyr Ser Ser Pro Glu Ile Thr
1 5 10 15
Val Cys Leu His Leu Phe Thr Ser Ser Gly Leu Lys Met Gln His Leu
20 25 30
Arg Ser His Cys His Leu Phe Arg Gln Ala Leu Phe Leu Cys Phe Ser
35 40 45
Asp Thr Thr Val Met Gly Phe Phe Leu Ser Tyr Trp Trp Gln Phe Ser
50 55 60
Val *
65

<210> 286
<211> 50
<212> PRT
<213> Homo sapiens

<400> 286
Met Trp Ser Val Thr Ser Thr Ile Phe Ile Cys Arg Leu Leu Ile Val
1 5 10 15
Arg Leu Leu Gly Asn Thr Ala Val Arg Thr Ser Val Val Phe Leu Pro
20 25 30
His Lys Ala Gly Arg His Trp Glu Lys Ser Thr Ser Leu Val Ser Gly

```
         35                  40                  45
Gly  *
 49
```

<210> 287
<211> 63
<212> PRT
<213> Homo sapiens

<400> 287

```
Met Trp Ala Asp Ser Ile Leu Ala Ser Leu Leu Leu Trp Pro His Gln
  1               5                  10                  15
Ser Leu Gln Leu Trp His His Pro His Leu Ala Asn Lys Asn Met Gly
            20                  25                  30
Val Pro Pro Pro Thr Thr Cys Lys Pro Trp Ser Thr Val Ala Gln Lys
        35                  40                  45
Phe Ala Asp Tyr Ile Pro Phe Met Thr Thr Trp Pro Pro Leu Gly
    50                  55                  60          63
```

<210> 288
<211> 170
<212> PRT
<213> Homo sapiens

<400> 288

```
Met Arg Leu Leu His Cys Lys Thr Leu His Ile Val Leu Phe Thr Leu
  1               5                  10                  15
Leu Tyr Lys Ile Leu Met Asp His Gln Asn Leu Ser Glu His Val Leu
            20                  25                  30
Cys Met Val Leu Tyr Leu Ile Glu Leu Gly Leu Glu Asn Ser Ala Glu
        35                  40                  45
Glu Glu Ser Asp Glu Glu Ala Ser Val Gly Gly Pro Glu Arg Cys His
    50                  55                  60
Asp Ser Trp Phe Pro Gly Ser Asn Leu Val Ser Asn Met Arg His Phe
 65                  70                  75                  80
Ile Asn Tyr Val Arg Val Arg Val Pro Glu Thr Ala Pro Glu Val Lys
                85                  90                  95
Arg Asp Ser Pro Ala Ser Thr Ser Ser Asp Asn Leu Gly Ser Leu Gln
            100                 105                 110
Asn Ser Gly Thr Ala Gln Val Phe Ser Leu Val Ala Glu Arg Arg Lys
        115                 120                 125
Lys Phe Gln Glu Ile Ile Asn Arg Ser Ser Ser Glu Ala Asn Gln Val
    130                 135                 140
Val Arg Pro Thr Thr Ser Ser Lys Trp Ser Ala Pro Gly Ser Ala Pro
145                 150                 155                 160
Gln Leu Thr Thr Ala Ile Phe Gly Asn  *
                165                 169
```

<210> 289
<211> 69
<212> PRT
<213> Homo sapiens

<400> 289

```
Met Lys Met Phe Gln Met Leu Leu Thr Ser Ser Phe Cys Ser Leu Ser
```

```
 1               5                  10                  15
His Leu Gln Ser Cys Gln His Ile Ser Phe Leu Ser Ile Ser Asn His
            20                  25                  30
Ser Lys Ile Phe Leu Tyr Leu Gln Pro Thr Cys Tyr Leu Tyr Leu Pro
        35                  40                  45
Pro Leu Pro Leu Phe Ser Arg Ser Trp His Trp Asn Leu Arg Val His
    50                  55                  60
Ile Cys Ser Pro  *
65          68
```

<210> 290
<211> 102
<212> PRT
<213> Homo sapiens

<400> 290

```
Met Cys Val Ala Ala Cys Phe Ser Leu Val Ala Trp Ser Ile Leu Gln
 1               5                  10                  15
Trp Gly Lys Arg Lys Tyr Pro Glu Gly Asn Ser Ser Trp Gln Ile Lys
            20                  25                  30
Glu Lys Val Trp Arg Phe Ser Thr Ala Phe Cys Ser Val Asn Glu Trp
        35                  40                  45
Lys Phe Ala Asp Ile Leu Ser Met Ala Asp His Leu Lys Lys Cys Ser
    50                  55                  60
Tyr Asn Val Val Glu Lys Arg Glu Glu Ala Ile Pro Leu Pro Cys Met
65                  70                  75                  80
Cys Val Thr Arg Glu Leu Thr Lys Glu Gly Arg Ser Leu Arg Ser Val
                85                  90                  95
Leu Lys Pro Val Leu  *
           100 101
```

<210> 291
<211> 68
<212> PRT
<213> Homo sapiens

<400> 291

```
Met Cys Thr Phe Arg Gly Leu Leu Thr Gly Leu Leu Thr Phe Pro Leu
 1               5                  10                  15
Phe Ser Pro Val Leu Tyr Phe Cys Asn Lys Phe Pro Asn Lys Thr Asn
            20                  25                  30
Met Phe Leu Leu Cys Phe Cys Lys Asn Tyr Phe Leu Ser Thr Val Phe
        35                  40                  45
Phe Ile Phe Leu Arg Gln Ser Phe Val Leu Val Ala Gln Thr Gly Val
    50                  55                  60
Gln Gly Val  *
65      67
```

<210> 292
<211> 105
<212> PRT
<213> Homo sapiens

<400> 292

```
Met Lys Gly Ile Leu Phe Phe Phe Phe Trp Lys Gly Val Tyr Phe Ser
```

```
 1               5                  10                  15
Pro Ser Leu Lys Pro Arg Gly Glu Ile Trp Val Asn Cys Pro Gln Pro
            20                  25                  30
Trp Gly Glu Gly Gly Pro Ile Gly Gly Lys Ile Lys Asn Gly Gly Val
        35                  40                  45
Phe Ser Gly Arg Glu Phe Phe Pro Thr Met Glu Lys Lys Lys Phe Pro
    50                  55                  60
Pro Arg Ala Lys Thr Lys Ile Asn Pro Pro Arg Lys Met Gly Ala Gln
65                  70                  75                  80
Arg Arg Pro Thr Pro Lys Trp Pro Thr Arg Gln Gly Pro Phe Asn Arg
                85                  90                  95
Ser Pro Lys Lys Gly Lys Arg Tyr Pro
            100                 105
```

<210> 293
<211> 95
<212> PRT
<213> Homo sapiens

<400> 293

```
Met Pro Trp Val Leu Gly Cys Thr Pro Phe Ile Ala Leu Ala Tyr Phe
 1               5                  10                  15
Phe Leu Trp Phe Leu Pro Pro Phe Thr Ser Leu Arg Gly Leu Trp Tyr
            20                  25                  30
Thr Thr Phe Tyr Cys Leu Phe Gln Ala Leu Ala Thr Val Pro Tyr Thr
        35                  40                  45
Ala Leu Thr Met Leu Leu Thr Pro Cys Pro Arg Glu Arg Asp Ser Ala
    50                  55                  60
Thr Ala Tyr Arg Met Thr Val Glu Met Ala Gly Thr Leu Met Gly Ala
65                  70                  75                  80
Thr Val His Gly Leu Ile Val Ser Gly Ala His Arg Pro His Arg
                85                  90                  95
```

<210> 294
<211> 52
<212> PRT
<213> Homo sapiens

<400> 294

```
Met Ala Val Glu Pro Leu Leu Ala His Phe Leu Arg Trp Ser Trp Leu
 1               5                  10                  15
Ser Ala Arg Asp Phe Tyr Ser Leu Gly Asn Val Asp Pro Ala Leu Trp
            20                  25                  30
Val Pro Cys Phe Phe Leu Leu Phe Leu Leu Ile Ile Thr Asp Asn Asn
        35                  40                  45
Asn Asp Ser  *
    50  51
```

<210> 295
<211> 52
<212> PRT
<213> Homo sapiens

<400> 295

```
Met Cys Ser Val Thr Cys Gly Val Leu Phe Ala Leu Ser Gly Leu Leu
```

```
  1               5                  10                  15
Leu Tyr Ser Ser Pro Ser Pro His Trp Asn Arg Pro Ser Arg Ile Ala
            20                  25                  30
Val Tyr Leu Met Cys Leu Thr Lys Tyr Cys Thr Gly Ser Ser Ala Ala
        35                  40                  45
Ser Cys Gln  *
     50  51
```

<210> 296
<211> 57
<212> PRT
<213> Homo sapiens

<400> 296

```
Met Pro Ala Cys Cys Tyr Arg Pro Cys Leu Leu Gln Pro Ile Ser Leu
  1               5                  10                  15
Leu Asn Ile Leu Leu Leu Leu Met Arg Lys Pro Ser Gln Glu Val Ile
            20                  25                  30
Asn Asp Thr Pro Lys Ala Gly Lys Trp Leu Ser Arg Tyr Leu Asp Ser
        35                  40                  45
Gly Leu Phe Tyr Ser Cys Ala Cys Gly
     50                  55      57
```

<210> 297
<211> 88
<212> PRT
<213> Homo sapiens

<400> 297

```
Met Trp Ser Trp His Val Gln Leu Gln Val Ser Ala Pro Leu His His
  1               5                  10                  15
Leu Leu Cys Leu His Phe Pro Pro Ala His Arg Ile Tyr Met Pro Phe
            20                  25                  30
Pro Ser Pro Lys Arg Ala Pro Ala Met Leu Asn Lys Gly Ile His Met
        35                  40                  45
Gln Gly Met Ser Ser Val Ser Trp Lys Gly Glu Ala Lys Phe Ser Phe
     50                  55                  60
His His Gln Arg Val Ala Phe Asn Ile Ile Tyr Thr Arg Gln Ala Phe
 65                  70                  75                  80
Ala Leu Leu Val Leu Leu Asn  *
                85      87
```

<210> 298
<211> 52
<212> PRT
<213> Homo sapiens

<400> 298

```
Met Gln Val Ser Gly Pro Arg Pro Gln Leu Phe Leu Pro Ser Val Phe
  1               5                  10                  15
Phe Val Leu Leu Phe Ser Tyr Thr Phe Thr Glu Thr Thr Gln Trp Thr
            20                  25                  30
Val Val Ile Leu Ala Leu Asn Ser Lys Leu Ser Phe Lys Glu Ile Glu
        35                  40                  45
Thr Ile Phe  *
```

50 51

<210> 299
<211> 62
<212> PRT
<213> Homo sapiens

<400> 299
Met Phe Met Pro Gly Thr Val Leu Arg Ile Leu Leu Ala Leu Pro Tyr
1 5 10 15
Leu Ile Leu Thr Lys Gln Val Gln Phe Phe Leu Phe Ser Asp Glu Ile
20 25 30
Met Ala Trp Lys Val Val Ala Pro Gly Leu Glu Leu Ser Ala Val Thr
35 40 45
Pro Asp Ser Thr Leu Phe Asn His Tyr Thr Ile Leu Ser *
50 55 60 61

<210> 300
<211> 83
<212> PRT
<213> Homo sapiens

<400> 300
Met Gln Val Lys Phe Ile Leu Lys Tyr Tyr Ile Ser Phe Leu Trp Lys
1 5 10 15
Thr Val Thr Ala Asn Gly Glu Thr Val Asn Met Ser Leu Leu Tyr Ile
20 25 30
Phe Thr Thr Met Glu Met Arg Lys Lys Ser Glu Val Gly Leu His Leu
35 40 45
Pro Ile Ser Ile Leu Lys Pro Phe Phe Thr Ile Val Leu Asp Glu Lys
50 55 60
Ile Val Thr Gly Gln Val Trp Gly Gly Glu Leu Phe Leu Leu Phe Cys
65 70 75 80
Lys Asp *
82

<210> 301
<211> 59
<212> PRT
<213> Homo sapiens

<400> 301
Met Pro Glu Leu Pro Thr Trp Val Leu Ala Leu Leu Pro His Pro Val
1 5 10 15
Val Leu Leu Ile Asp Ser Gly Glu Leu Glu Ala Phe Glu Gln Ile Cys
20 25 30
Arg Ser Thr Leu Lys Ala Val Trp His Ser Val His Gly Ala Met Ser
35 40 45
Val Cys Phe Ile Cys Phe Thr Phe Cys His *
50 55 58

<210> 302
<211> 82

<212> PRT
<213> Homo sapiens

<400> 302

```
Met Arg Ile Val Arg Arg Met Cys Met Trp Ser Ala Gly Pro Ala Pro
 1               5                  10                  15
Ala Thr Val Cys Ala Val Met Val Ala Ala Pro Lys Ser Pro Gln Ser
            20                  25                  30
Pro Pro Arg Trp Ala Cys Val Tyr Ser Leu Ile Gly Cys His Ser Ser
        35                  40                  45
Asp Pro Phe Ser Val Tyr Phe Ser Gly Ile Ser Trp Arg Asp Ile Ser
    50                  55                  60
Leu Ser Leu Tyr Ser Met Ala Gln Glu Ser Gln Asn Gln Ser Ile Leu
65                  70                  75                  80
Lys  *
 81
```

<210> 303
<211> 83
<212> PRT
<213> Homo sapiens

<400> 303

```
Met Cys Pro Leu Leu Val Tyr Lys Ile Ile Leu Val Phe Ala Ala Met
 1               5                  10                  15
Phe Phe Phe Ser Gln Gly Ser Gln Val Glu Ile Arg Ser His Glu Gly
            20                  25                  30
Glu His Cys Val Gly Thr Val His Leu Leu Ser His Phe Leu Tyr Ser
        35                  40                  45
Lys Asn Asn Pro Val Phe Tyr Lys Gly Asn Thr Ser Phe Ile Phe Glu
    50                  55                  60
Thr Met Glu Glu Asp Ser Leu Ser Ser Leu Ala Glu Arg Ser Gly Ser
65                  70                  75                  80
Cys Met  *
     82
```

<210> 304
<211> 118
<212> PRT
<213> Homo sapiens

<400> 304

```
Met Pro Gln Phe Pro Val Ala Phe Gly Ile Met Phe Thr Tyr Phe Thr
 1               5                  10                  15
Leu Ala His Lys Val Leu His Ser Gln Ala Ser Ala Cys Leu Phe Ser
            20                  25                  30
Ile Ile Cys Phe Phe Pro Thr Cys Thr Leu His Phe Ser Gln Val Gly
        35                  40                  45
Ser His Ala Ala Pro Trp Met Gly His Asp Ala Leu Cys Leu Arg Val
    50                  55                  60
Phe Leu Tyr Arg Leu Pro Cys Glu Lys Pro Ser Pro Ser Ala His Met
65                  70                  75                  80
Val Thr Gly Ser Val Leu Glu Gly Pro Leu Cys Ala Leu Ala Leu Ser
                85                  90                  95
Ser Phe Pro Pro Gly Ala Thr Leu His Leu Ser Cys Leu Ser Leu Lys
            100                 105                 110
Arg Ala Val Phe Phe Tyr
```

115 118

<210> 305
<211> 73
<212> PRT
<213> Homo sapiens

<400> 305

```
Met Ser Trp Arg Thr Arg Ser Met His Thr His Ile Ser Val Ser Phe
 1               5                  10                  15
Lys Gly Lys Ile Arg Pro Thr Ser Ala Tyr Leu Leu Leu Phe Leu Phe
            20                  25                  30
Phe Phe Cys Tyr Gly Val Ser Leu Cys Cys Pro Gly Trp Ser Glu Val
        35                  40                  45
Val Ala Arg Ser Arg His Leu Ala Ser Ser Ala Ser Arg Val His Ala
    50                  55                  60
Ile Leu Leu Pro Gln Asn Pro Glu *
65                  70      72
```

<210> 306
<211> 47
<212> PRT
<213> Homo sapiens

<400> 306

```
Met Leu Ser Thr Leu Ser Ile Gly Thr Leu Ser Met Leu Ile Ile Val
 1               5                  10                  15
Val Ser Asp Ser Trp Ser Tyr Ser Ser Asn Ser Pro Ala Met Phe Gly
            20                  25                  30
Ser Asp Ala Gly Phe Ile Pro Ser Asn Cys Ile Phe Ala Phe  *
        35                  40                  45  46
```

<210> 307
<211> 94
<212> PRT
<213> Homo sapiens

<400> 307

```
Met Asp Pro Pro Cys Pro Trp Leu His Pro Ala Ala Trp Pro Leu Gln
 1               5                  10                  15
Thr Pro Leu Ala Leu Pro Leu Leu Gly Thr Gly Ser Ser Pro Met Pro
            20                  25                  30
Ile Phe Arg Trp Arg Pro Pro Val His Leu Leu Ser Met Ala Gln Gly
        35                  40                  45
Pro Ser Phe Leu Ala Gly Ala Ala Arg Gly Asp Lys Ala Lys Gly Ala
    50                  55                  60
Pro Arg Arg His Gly Ala Asn Phe Ala Leu Thr Arg Trp Ala Tyr Pro
65                  70                  75                  80
Ile Arg Ala Leu Asn Leu Leu Gly Gly Arg Gln Thr Trp  *
                85                  90          93
```

<210> 308
<211> 54

<212> PRT
<213> Homo sapiens

<400> 308

```
Met Ile Leu Val Ser Leu Leu Ile Leu Ile Val Glu Pro Leu Phe Ala
 1               5                  10                  15
Ser Leu Thr Pro Leu Ser Leu Cys Phe Glu Cys Val Val Phe Leu Asn
            20                  25                  30
Val Gly Gln His Leu Thr Asp Gln Thr Phe Ser Leu Asn Gly Leu Leu
        35                  40                  45
Phe Leu Ser Asn Ser  *
    50          53
```

<210> 309
<211> 68
<212> PRT
<213> Homo sapiens

<400> 309

```
Met Ser Pro Leu Leu Pro Leu Ser Tyr Lys Leu Val Leu Cys Phe Pro
 1               5                  10                  15
Thr Pro Asn Gly Val Val Thr His Gly Glu Gln Asn Ala Ser Ser Thr
            20                  25                  30
Asp Ile Glu His Gly Leu Lys Thr Ile Leu Ile Lys Pro Pro Ala Arg
        35                  40                  45
Ile Leu Lys Arg Lys Thr Glu Gly Glu Glu Ser Asn Arg Leu Thr Leu
    50                  55                  60
Pro Thr Thr  *
65      67
```

<210> 310
<211> 87
<212> PRT
<213> Homo sapiens

<400> 310

```
Met Gly Pro Val Ser Gly Cys Trp His Met Ser Leu Cys Leu Arg Val
 1               5                  10                  15
Tyr Leu Ala Leu Asp Pro Ala His Gln Glu Leu Met Pro Pro Gly Ser
            20                  25                  30
Ser Leu Gln Pro Ile Thr Leu Gly Ile Gly Ile Glu Ile Leu Gln Pro
        35                  40                  45
Pro Thr Leu Glu Val Gly Asn Ser Glu Ala Leu Ser Val Pro Ser Arg
    50                  55                  60
Arg Thr Pro Arg Arg Thr Glu Leu Pro Trp Pro Thr Val Leu Thr Gly
65                  70                  75                  80
Phe Leu Ile Asn Thr Leu  *
                85  86
```

<210> 311
<211> 53
<212> PRT
<213> Homo sapiens

<400> 311
Met Leu Thr Cys Val Pro Glu Arg Leu Phe Gln Cys His His Leu Ile
1 5 10 15
Arg Met Thr Cys Leu Phe Met Ile Leu Glu Phe Arg Leu Phe Lys Tyr
20 25 30
Asp Ser Asn Leu Cys Ser His Val Ile Ile Asn His Pro Gln Val Gln
35 40 45
Gly Arg Gln Arg *
50 52

<210> 312
<211> 51
<212> PRT
<213> Homo sapiens

<400> 312
Met Trp Gly Ala Pro Ala Leu Gln Cys Ile Val Phe Phe Arg Trp Thr
1 5 10 15
Arg Ser Lys Cys Leu Pro Asp Thr Gly Asn Val Cys Thr Lys Thr Gln
20 25 30
Arg Lys Lys Ala Ala Gly Arg Leu Gly Val Ala Gly Gly Ile Ala Leu
35 40 45
Gly Leu *
50

<210> 313
<211> 61
<212> PRT
<213> Homo sapiens

<400> 313
Met Ser Leu Arg Ile Arg Ala Ala Arg Asn Trp Ala Arg Asp Val Gln
1 5 10 15
Lys Leu Trp Thr Ile Val Val Leu Leu Val Leu Ile Leu Ile Arg Ser
20 25 30
Ala Val Asn Leu Leu Ile Asn Ser Arg Thr Glu Asp Lys Ser Leu Gln
35 40 45
Leu Val Leu Tyr Gln Ser Val Ile Ile Cys Phe Pro *
50 55 60

<210> 314
<211> 59
<212> PRT
<213> Homo sapiens

<400> 314
Met Val Trp His Val Arg Lys Ser Ser Phe Val Trp Leu Leu Gln Leu
1 5 10 15
Phe Ser Phe Ile Ser Cys His Ser Val Ile Ser Val Ser Pro Val His
20 25 30
Val Pro Trp Thr Gln Cys Ala Val Ile Pro Pro Tyr Thr Ser Cys Pro
35 40 45
Lys Leu Phe Ala Ile Gln Gly Gly Arg Phe *
50 55 58

<210> 315
<211> 81
<212> PRT
<213> Homo sapiens

<400> 315
Met Gly Gly Phe Lys Val Asn Leu Phe Val Lys Val Lys Ala Glu Gly
1 5 10 15
Ser Pro Leu Cys Trp Leu Lys Leu Ala Cys Leu Gly Ala Trp Leu Leu
20 25 30
Ser Leu Leu Ile Ser Gln Lys Ser Asp Glu Gln Leu Cys Phe Gly Leu
35 40 45
Leu Thr Trp Asn Phe Ser Thr Ser Asp Ser Ile Leu Val Trp Phe Val
50 55 60
Gly Pro Arg Ala Gly Thr Gln Ser Lys Pro Val Val Ser Tyr Lys Phe
65 70 75 80
*

<210> 316
<211> 52
<212> PRT
<213> Homo sapiens

<400> 316
Met Glu Thr Ser Ser Ala Phe Thr Asn Pro Leu Leu Val Cys Phe Leu
1 5 10 15
Ala Leu Leu His Ser Val Met Asn Ile Thr Tyr Thr Pro Pro Lys Lys
20 25 30
Lys Asn Glu Asn Cys Ser Lys Pro Leu Ile Leu Thr Ser Ser Leu Gly
35 40 45
Thr Val Gln *
50 51

<210> 317
<211> 73
<212> PRT
<213> Homo sapiens

<400> 317
Met Cys Ala Phe Leu Leu Pro Phe Lys Leu Leu Phe Phe Leu Glu Ile
1 5 10 15
Ser Leu Ala Met Lys Ser His Phe Pro Phe Thr Leu Leu Ile Leu Ser
20 25 30
Arg Val Leu Leu Lys Lys Thr Leu Tyr Val Leu Lys Leu Gly Trp Leu
35 40 45
Ile Thr Ile Pro Ser Asn Asp Leu Thr Ser Val Phe Thr Leu Met Ile
50 55 60
His Arg Gln Asn Gln Lys His Phe *
65 70 72

<210> 318
<211> 70

<212> PRT
<213> Homo sapiens

<400> 318
Met Pro Thr Leu Leu Gln Val Met Ser Trp Met Leu Ser Phe Gly Thr
1 5 10 15
Gln Thr Leu Gln Leu Glu Ser Cys Thr Cys Ala Leu His Ile Val Gly
20 25 30
Ala Trp Lys Val Pro Tyr Pro Leu Phe Ser Arg Val Leu Ile Cys Gln
35 40 45
Val Lys Ile Leu Ser Thr Ser Ile Ser Gln Glu Lys Val Phe Arg Thr
50 55 60
Glu Ser Arg Thr Glu *
65 69

<210> 319
<211> 46
<212> PRT
<213> Homo sapiens

<400> 319
Met Cys Arg Ala Leu Leu Leu Leu Leu Cys Ser Pro Asn Ser Ser Phe
1 5 10 15
Gln Trp Leu Pro Leu Pro Val His Pro His Thr Thr Ile Arg Tyr Arg
20 25 30
Ser Tyr Asn Met Val Pro Val Lys Leu Thr Asn Val Gln *
35 40 45

<210> 320
<211> 56
<212> PRT
<213> Homo sapiens

<400> 320
Met Leu Phe Leu His Ile Ala Glu Cys Ser Phe Leu Arg Leu Lys Val
1 5 10 15
Ala Phe Pro Ser Ser Leu Asn Phe Gln Pro Leu Ala Gln Phe Leu Ala
20 25 30
His Ile Leu Glu Val Phe Tyr Lys Cys Leu Trp Lys Lys Gly Val Gln
35 40 45
Val Phe Asn Phe Leu Ala Asn *
50 55

<210> 321
<211> 58
<212> PRT
<213> Homo sapiens

<400> 321
Met Pro Leu Asp Thr Asp Ala Ile Leu His Arg Thr Ala Glu Trp Tyr
1 5 10 15
Val Leu Cys Leu Ile Thr Cys Ile Phe Met Tyr Val Leu Tyr Val Pro
20 25 30
Tyr Leu Arg Ser Leu Ile Leu Leu Glu Tyr Leu His Leu Leu Pro Phe

```
         35                  40                  45
Glu Ile Leu Ile Gln Ala Asn Ala Gly  *
     50                  55      57
```

<210> 322
<211> 50
<212> PRT
<213> Homo sapiens

<400> 322

```
Met Ala Thr Arg Ser Lys Gly Ala Phe Ile Asn Cys Tyr Ile Ile Leu
  1               5                  10                  15
Leu Leu Thr Phe Leu Met Ile Arg Thr Phe Tyr Asn Leu Met Glu Tyr
             20                  25                  30
Tyr Cys Pro Thr Leu Leu Ile Arg Lys Leu Met Ser Asn Thr Lys Ile
         35                  40                  45
Leu  *
 49
```

<210> 323
<211> 53
<212> PRT
<213> Homo sapiens

<400> 323

```
Met Leu Thr Cys Val Pro Glu Arg Leu Phe Gln Cys His His Leu Ile
  1               5                  10                  15
Arg Met Thr Cys Leu Phe Met Ile Leu Glu Phe Arg Leu Phe Lys Tyr
             20                  25                  30
Asp Ser Asn Leu Cys Ser His Val Ile Ile Asn His Pro Gln Val Gln
         35                  40                  45
Gly Arg Gln Arg  *
     50      52
```

<210> 324
<211> 75
<212> PRT
<213> Homo sapiens

<400> 324

```
Met Asp Phe Leu Ser Arg Leu Met Leu Leu Arg Met Cys Lys Cys Val
  1               5                  10                  15
Thr Ala Thr Tyr Gln Tyr Ile Arg Arg Ser Leu Phe Leu Asn Leu Val
             20                  25                  30
Pro Leu Leu Gln Thr Leu Ser Ile Leu Thr Ala His Ser Val Leu Leu
         35                  40                  45
Arg Pro Ala Leu Ser Ser Leu Val Lys Met Glu Asp Ser Gln Ala Leu
     50                  55                  60
Ser Leu Ser Leu Glu Pro Glu Ser Ala Phe  *
 65                  70                  74
```

<210> 325
<211> 61

<212> PRT
<213> Homo sapiens

<400> 325

```
Met Phe Cys Leu Trp Asn Gln Trp Val Val Thr Ala Gln Arg Leu Leu
 1               5                  10                  15
Val Ser Trp Leu Ser His Ala Gln Arg Gln Pro Cys Pro Leu Ser Leu
            20                  25                  30
Phe Cys Gly Arg Arg Asn Pro Leu Ala Trp Thr Ile Phe Gly Trp Lys
        35                  40                  45
His Gln Pro Leu Thr Ser Asp Cys His Phe Gln Met  *
    50                  55                  60
```

<210> 326
<211> 59
<212> PRT
<213> Homo sapiens

<400> 326

```
Met Thr Thr Ser Ser Leu Val Leu Pro Pro Leu Phe Val Leu Lys Cys
 1               5                  10                  15
Gln Arg Phe Tyr Pro Pro Leu Tyr Leu His Pro Tyr Ser Ile Cys Gln
            20                  25                  30
His Val Ser Ile Leu Val Lys Ile Val Trp Thr Trp Gly Ser Glu Val
        35                  40                  45
Pro Thr Leu Gly Thr Ile Glu Ile Gly Thr  *
    50                  55          58
```

<210> 327
<211> 73
<212> PRT
<213> Homo sapiens

<400> 327

```
Met Leu Ala Trp Arg Leu Leu Cys Pro Trp Gly Pro Gly Leu Pro Thr
 1               5                  10                  15
Thr Thr Ala Arg Ser Gly Glu Arg Thr Glu Arg Arg Glu Arg Val Arg
            20                  25                  30
Thr Ala Ser Pro Arg Lys Ile Leu Phe Lys Thr Gln Pro Pro Arg Gly
        35                  40                  45
Ser Ser Thr Asp Arg Cys Pro Trp Gly Arg Gln Cys Leu His Gly Thr
    50                  55                  60
Gly Thr Cys His Met Pro Asn Arg  *
65                  70      72
```

<210> 328
<211> 47
<212> PRT
<213> Homo sapiens

<221> misc_feature
<222> (1)...(47)
<223> Xaa = any amino acid or nothing

<400> 328
Met Thr Ser Arg Pro His Phe Phe Arg Tyr Leu Cys Ser Leu Pro Pro
1 5 10 15
Leu Leu Phe Pro Leu Leu Xaa Gln Ser Gln Leu Leu Pro Gly Ser Pro
20 25 30
Leu Pro Ile Ala Leu Gln Ser Arg Val Gly Ser Leu Leu Ala *
35 40 45 46

<210> 329
<211> 96
<212> PRT
<213> Homo sapiens

<400> 329
Met Leu Pro Ser Phe Leu Pro Gln Ser Leu Gly Asn Leu Ile His Thr
1 5 10 15
Leu Gly Phe Leu Leu Ile Ile His Lys Tyr Met Ser Ala Phe Lys Asn
20 25 30
Arg Thr Asp Glu Phe Met Asn Met Gly Met Gln Pro Tyr Ile Lys Ser
35 40 45
Pro Tyr Arg Leu Ser Met Ser Gln Ile Ser Leu Lys Phe Asp Leu Ser
50 55 60
Gln Thr Asp Leu Ile Leu Pro His Lys Phe Tyr Ser Pro Ser Ser Phe
65 70 75 80
Pro Thr Val Met Leu Phe Tyr Ser Phe Gly Arg Leu Ser His Lys Pro
85 90 95 96

<210> 330
<211> 50
<212> PRT
<213> Homo sapiens

<400> 330
Met Gln Ala Trp Arg Ser Phe Val Met Gly Val Glu Val Leu Met Tyr
1 5 10 15
Ile Val Ala Val Arg Cys Arg Ala Val Phe Ala Thr Ser Leu Trp Gln
20 25 30
Pro Trp Cys Tyr Thr Arg Ala Gly Gly Gln Phe Asn Val Ser Gln Ala
35 40 45
Arg *
49

<210> 331
<211> 78
<212> PRT
<213> Homo sapiens

<400> 331
Met Val Val Leu Tyr Ile Val Arg Ala Tyr Asn His Tyr Ile Leu Cys
1 5 10 15
Cys Leu Ser Ser Ser Leu Tyr Leu Val Leu Ile Leu Leu Val Thr Val
20 25 30
Tyr Leu Met Leu Thr Thr Ser Ser Tyr Asn Asp Val Ser Leu Val Ile

```
         35                  40                  45
Trp Ile Ala Ser Ser Phe Ala Ser Ser Lys Phe Phe Arg Lys Gly Leu
     50                  55                  60
Arg Glu Tyr Ser Tyr Phe Met Asn Phe Leu Ala Arg Ser  *
 65                  70                  75      77
```

<210> 332
<211> 93
<212> PRT
<213> Homo sapiens

<400> 332

```
Met Leu Glu Trp Pro Leu Leu Gly Gln Ile Leu Pro Met Ile Ile Pro
  1               5                  10                  15
Leu Pro Pro Leu Pro Ala Leu Val Val Trp Pro Ile Gly Leu Thr His
            20                  25                  30
Cys Pro Trp Pro Ser Pro Phe Val Pro Ala Ser Leu Asp Gly Phe Tyr
         35                  40                  45
Asn Ser Arg Ser Leu Asp Gly Lys Ser Pro Pro Leu Arg Pro Glu Lys
     50                  55                  60
Trp Ser Pro Trp Ser Trp Phe Ser Arg Leu Pro Gly His Gly Leu Pro
 65                  70                  75                  80
Lys Glu Gly Gly Gln Arg Lys Arg Glu Val Trp Ala  *
                85                  90      92
```

<210> 333
<211> 66
<212> PRT
<213> Homo sapiens

<400> 333

```
Met Met Leu Trp Gln Val Tyr Pro Gly Pro Ser Ala Ala Val Leu Cys
  1               5                  10                  15
Leu Phe Leu His Pro Pro Trp Ser Arg Ser Thr Ala Val Glu Arg Glu
            20                  25                  30
Lys Arg Gln Lys Asp Gly Arg Gly Gln Arg Met Leu Leu Pro Gln Pro
         35                  40                  45
Gln Cys Leu Met Ser Ser Cys Cys Leu Val Asp Val Gln Ser Leu Thr
     50                  55                  60
Gly  *
 65
```

<210> 334
<211> 61
<212> PRT
<213> Homo sapiens

<400> 334

```
Met Phe Phe Cys Leu Thr Arg Gln Ser Leu Leu Cys Thr Leu Leu Leu
  1               5                  10                  15
Met Leu Lys Arg Cys Ile Phe Phe Ser Tyr Cys Val Ile Cys Arg Ala
            20                  25                  30
Lys Ser Phe Glu Leu Phe Thr Ser Glu Ile Thr Phe Pro Asp Lys Arg
         35                  40                  45
Ala Lys Gln Cys Leu Phe Lys Leu Phe Ser Gly Thr  *
```

```
     50                  55                  60
```

<210> 335
<211> 53
<212> PRT
<213> Homo sapiens

<400> 335

```
Met Gly Asn Thr Asp Ile Leu Leu Leu Leu Ser Leu Phe Cys Phe Ser
  1               5                  10                  15
Tyr Glu Leu Val Ala Gly Lys Thr Lys Ala Gln Phe Gly Val Pro Phe
             20                  25                  30
Ala Glu Phe Ser Val Phe Leu Ile Leu Glu Asn Val Thr Ala Cys Arg
         35                  40                  45
Phe Leu Tyr Ile  *
     50      52
```

<210> 336
<211> 69
<212> PRT
<213> Homo sapiens

<400> 336

```
Met Ser Ile Leu Val Val Ser Ala Phe Leu Ala Asn Leu Trp Leu Leu
  1               5                  10                  15
Met Thr Ile Ser Thr Ser Gln Met Leu Asn Met Thr Lys Ile Thr Tyr
             20                  25                  30
Leu Val Leu Phe Leu His Leu Ser Ala Leu Arg Ile Gly Ser Thr Pro
         35                  40                  45
His Ser Phe Leu Leu Lys Ser Tyr His Leu Gly Thr His Phe Ser Leu
     50                  55                  60
Phe His Met Asn Ser
 65              69
```

<210> 337
<211> 79
<212> PRT
<213> Homo sapiens

<400> 337

```
Met Phe Val Phe Val Ser Ile His Thr Glu Leu Val Pro Ile Leu Arg
  1               5                  10                  15
Pro Leu Cys Leu Leu Tyr Cys Cys Pro Asp Cys Ser Val Pro Arg Pro
             20                  25                  30
Leu Tyr Ser Leu Lys Tyr Leu Leu Leu Ile Asn Asp Phe Pro Glu Leu
         35                  40                  45
Gln Ile His Met Ser Ser Phe Ser Gln Ser Leu His Tyr Ile Ile Leu
     50                  55                  60
Ser Tyr Phe Phe His Ser Ile Cys His Ile Leu Leu Val Leu  *
 65                  70                  75          78
```

<210> 338
<211> 53

<212> PRT
<213> Homo sapiens

<400> 338
Met Met Arg Val Ile Ile Leu Ile Trp Phe Arg Ile Ser Lys Gly Thr
1 5 10 15
Phe Gln His Ser Thr Thr Lys Cys Asp Val Cys Phe Arg Val Phe Leu
20 25 30
Leu Ser Asn Cys Ser Phe Leu Ser Leu Asn Tyr Lys Leu Thr Ser Asp
35 40 45
Phe Ile Ile Tyr *
50 52

<210> 339
<211> 63
<212> PRT
<213> Homo sapiens

<400> 339
Met Gln Met Val Val Pro Arg Leu Leu Ser Val Pro Gln Leu Leu Asn
1 5 10 15
Thr Ala Pro Leu Phe Leu Pro Trp Glu Lys Thr Val Lys Thr Gln Tyr
20 25 30
Ser Gly Ile Ile Phe Lys Phe Lys Ser Arg Ile Glu Thr Ala Glu Lys
35 40 45
Ser Ile Gly Asp Thr Lys Glu Arg Ile Gln Pro Ser Gln Ile *
50 55 60 62

<210> 340
<211> 60
<212> PRT
<213> Homo sapiens

<400> 340
Met Leu Leu Pro Val Phe Leu Leu Tyr Leu Ser Gln Asp Leu Ala Asp
1 5 10 15
Ser Arg Ala Pro Ala His Cys Ser Val Asn Thr Asp Leu His Leu Lys
20 25 30
Trp Gly Ser Leu Cys Val Leu Ser His Phe Gln Val Asp Leu Pro Val
35 40 45
Asn Pro Ile Cys Glu His Ile Cys Arg Cys Pro *
50 55 59

<210> 341
<211> 46
<212> PRT
<213> Homo sapiens

<400> 341
Met Asn Leu Glu His Val Ile Val Ser Leu Leu Thr Phe Tyr Arg Val
1 5 10 15
Leu Leu Tyr Lys Glu Ile Ile Gly Leu His His Cys Phe Gln His Phe
20 25 30
His Val Asn Ala Phe Leu Leu Ser Pro Leu Pro Pro Ser *

35 40 45

<210> 342
<211> 48
<212> PRT
<213> Homo sapiens

<400> 342
Met Ser Ala Leu Trp Leu Val Tyr Val Ser Cys Gly Cys Ser Cys Ala
1 5 10 15
Leu Cys Val Leu Val His Ala Gln Thr Cys Thr Ser Phe Ser Ile Leu
20 25 30
Val His Ala Gln Ala His Val Leu Ser Cys Asp His Ile His Thr Ser
35 40 45 48

<210> 343
<211> 51
<212> PRT
<213> Homo sapiens

<400> 343
Met Leu Ile Gln Phe Thr Phe Phe Phe Pro Ile Ser Gln Arg Phe Trp
1 5 10 15
Phe Cys Leu Leu Phe Phe Phe Pro Gln Thr Phe Lys Cys Met Lys Phe
20 25 30
Tyr Ser Leu Ile Glu Val Arg Glu Gly Val Cys Ile His Gln Arg Ser
35 40 45
Gln Phe Val
50 51

<210> 344
<211> 93
<212> PRT
<213> Homo sapiens

<400> 344
Met Ile Glu Leu Ala Trp Lys Phe Ile Met His Ile Asn Ala Leu Leu
1 5 10 15
Ser Phe Gly Thr Thr Ser Leu Lys Ile Lys Phe Ala His Val Phe Ser
20 25 30
Phe Leu Gln Leu Cys Ile Met Asn Lys Ile Thr Leu Ile Ser Leu Asn
35 40 45
Thr Pro Pro Phe Gln Glu Glu Phe Thr Glu Leu Gly Ser Phe Val Asn
50 55 60
His Cys Glu Leu Ile Asn Lys Asn Leu Ser His Asn Leu Pro Phe Phe
65 70 75 80
Pro Trp Leu Ile Ile Leu Leu Lys Val Ile Arg Tyr *
85 90 92

<210> 345
<211> 70

<212> PRT
<213> Homo sapiens

<400> 345
Met Pro Asn Arg Ile Glu Leu Ser Ala Leu Phe Arg Val Pro Phe His
1 5 10 15
Val Asp Trp Leu Leu Phe Ala Phe Val Phe Phe Pro Leu Cys Tyr Val
20 25 30
Leu Gly Phe Cys Leu Pro Ser Thr Val Lys Thr His Ser Leu Ile Val
35 40 45
Gly Ala Val Phe Gln Ser Glu Ile Val Thr Asn Val Cys Gln Ile Tyr
50 55 60
Lys Cys Lys Thr Thr *
65 69

<210> 346
<211> 75
<212> PRT
<213> Homo sapiens

<400> 346
Met Arg Ala His Pro His Thr Pro Arg Ser Ala Ala His Ile Ile Leu
1 5 10 15
Ser Thr Trp Leu Ala Leu Gly Pro Tyr Leu Tyr Arg Thr Arg Asn Cys
20 25 30
Gly Pro Leu Thr Glu Glu Pro Ser Leu Ala Pro Ser Thr Gln Gly Ala
35 40 45
Arg Arg Thr Asn Thr Thr Pro Glu Leu Ala Ala Arg Pro Val Ala Pro
50 55 60
Cys Leu His Ile Ser Tyr Thr Phe Thr Arg Gly
65 70 75

<210> 347
<211> 72
<212> PRT
<213> Homo sapiens

<400> 347
Met Pro Leu Leu Tyr Ile Ile Cys Leu Arg Gln Leu Val Leu Phe His
1 5 10 15
Ser Lys Cys His Ser Gln His Ser Cys Arg Ala Gly Gly Ile Gln Tyr
20 25 30
Ser Met His Val Ser Leu Phe Leu Ser Ser Pro Ile Asn Tyr Asp Asn
35 40 45
Gly Phe Leu Val Ser Pro Thr Phe Pro Leu His Val Lys Leu Ser Phe
50 55 60
Leu Lys Tyr Ser Phe Lys Cys Ile
65 70 72

<210> 348
<211> 60
<212> PRT
<213> Homo sapiens

<400> 348

```
Met Gly Pro Lys Arg Gly Leu Phe Phe Phe Ile Phe Phe Leu Asp Thr
 1               5                  10                  15
Glu Pro Ser Val Leu Gly Gly Gly Gly Gly Gly Gln Tyr Gly Leu Thr
            20                  25                  30
Arg Thr His Leu Trp Arg Gln Gly Ala Ser Tyr Leu Thr Leu Leu Arg
        35                  40                  45
Asn Gly Thr Gln Ser Gly Pro Leu Ala His Leu  *
    50                  55              59
```

<210> 349
<211> 751
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (116)..(319)

<220>
<221> misc_feature
<222> (1)...(751)
<223> n = a,t,c or g

<400> 349

```
atcatgatgc atatacctat ctattcgatg atgaagatac cccaccaaac ccaaaaaaag     60

agatctctcg aggatccgaa ttcgcggccg cgtcgactgc actattaagt agacc atg     118
                                                              Met
                                                               1

tct ccc ctt ctc cca ctg tcc tat aaa ttg gtc ctg tgt ttc cca acc      166
Ser Pro Leu Leu Pro Leu Ser Tyr Lys Leu Val Leu Cys Phe Pro Thr
            5                  10                  15

cca aat ggg gta gta act cat gga gaa caa aat gcc agt agt aca gat      214
Pro Asn Gly Val Val Thr His Gly Glu Gln Asn Ala Ser Ser Thr Asp
        20                  25                  30

ata gaa cat ggg tta aaa act att ctg atc aaa ccc cca gca aga atc      262
Ile Glu His Gly Leu Lys Thr Ile Leu Ile Lys Pro Pro Ala Arg Ile
    35                  40                  45

ttg aag agg aag aca gag ggg gaa gaa agc aat agg ctt act ctc cct      310
Leu Lys Arg Lys Thr Glu Gly Glu Glu Ser Asn Arg Leu Thr Leu Pro
 50                  55                  60                  65

aca act taa ttccatg ctctaattca gagcagtata attaaccccg agtttcactg      366
Thr Thr  *

ctaacactga agtttttcag cctgatgaga atatcattta gtatcctgca atcgccaata    426

atagcaatat gacacaatgg tgcccattga ggctctaatc agagcacata gtcatttaca    486

tagaacattt tctgtgactc tctgtaatac aacgccacaa agactcaatt gagacacaag    546

tctccctttc ccataancta nnnaaaaaaa aaaaccaaga aaaaaggggg ggggggggga    606

aaaaggaaaa gagggacccc ccttttttttt ttcaaggaga ccaaaatggt tccccaccat    666
```

```
tattccgggg ggttcacaag cccggggttt ttctttaaaa aaagttattt ccggtggcgg    726

aaaaaaaata aaaacctttc tcttt                                          751
```

<210> 350
<211> 1122
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (211)..(582)

<400> 350

```
tcccggttcg acgattcgtt gtcctataaa atcccaaaat gagtgttgct tcatgacagt     60

ttctctttct gtttttattg atgagttaat ctttttctcc gatgtgggca gggatttaga    120

catagtggac tagtggctgt tgcctgtttc tagtgaggac ttctgctgcc ctttctgttt    180

tgagttccca ttctgcaaac ccccccctcac     atg cgg ccc ctg ccg ttc cct     231
                                      Met Arg Pro Leu Pro Phe Pro
                                        1               5

acc aaa tat cta ggt aag ttc agt tgg aag tgg ttt ttt ctt tct ctg      279
Thr Lys Tyr Leu Gly Lys Phe Ser Trp Lys Trp Phe Phe Leu Ser Leu
         10                  15                  20

aat gct ggt gac tat cat aga gaa ttt agc aca gtc tac ttt gaa ttt      327
Asn Ala Gly Asp Tyr His Arg Glu Phe Ser Thr Val Tyr Phe Glu Phe
     25                  30                  35

gaa agc ttc ggg cat tgg tct ttt ctt cag tgc ttc atg tcc tgg aga      375
Glu Ser Phe Gly His Trp Ser Phe Leu Gln Cys Phe Met Ser Trp Arg
 40                  45                  50                  55

aca aga agc atg cac aca cac att tca gtt tct ttt aaa gga aaa ata      423
Thr Arg Ser Met His Thr His Ile Ser Val Ser Phe Lys Gly Lys Ile
                 60                  65                  70

aga ccc aca tct gct tat ctg ctt ctt ttt ctt ttt ttc ttt tgt tat      471
Arg Pro Thr Ser Ala Tyr Leu Leu Leu Phe Leu Phe Phe Phe Cys Tyr
             75                  80                  85

gga gtc tcg ctt tgt tgc cca ggc tgg agt gaa gtg gtg gcg cga tct      519
Gly Val Ser Leu Cys Cys Pro Gly Trp Ser Glu Val Val Ala Arg Ser
         90                  95                 100

cgg cac ttg gca agc tcc gcc tcc cgg gtt cac gcc att ctc ctg cct      567
Arg His Leu Ala Ser Ser Ala Ser Arg Val His Ala Ile Leu Leu Pro
    105                 110                 115

cag aat ccg gag tag ctgggactac aggcgcccgc caccacgctc gggtaatttt      622
Gln Asn Pro Glu  *
120

ttgtattttt agtagagacg gggtttcacc gtgttagcca ggatggtctc gatctcctga    682

cctcatgatc gcccgcctcg gcctccccaa gtgccaggat tacaggcgtg agccaccgtg    742
```

```
cccagccact taactgtttt ttaaacaaac atgaacttaa aattatatca aacagaccat    802

tattgcgccg gtctacagac ccagcttact agcgggctgc acgcatagtc ttctttacgt    862

cactaaatca atcccggccg ccgtttacac agcggactgg aacccccggg ctcaaaaaac    922

gaattctatc ggggctagcc cgctccgttc ttacaaaacg aatccccccc cccgcgtgcg    982

ccggcaggag ccataccgac ctgcatatcc gggccccccc ggctcggctc cggaccaacc   1042

cttgcttgag cggccccgcg ggccgtctac tccggtaaaa ctcatcaccc ccgccctgcc   1102

cgctctggcc gggtaagacc                                              1122
```

<210> 351
<211> 1282
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (133)..(1161)

<400> 351

```
taccggtccg gaattcccgg gtcgacgatt tcgtgcagga agtgactgcg ggagtggagc     60

cggcgagaga gtggcagcgg gggctgatgg aagtgcagtg ggggctggag agggcaccct    120

actgtatcca gc    atg ctc caa ggc cac agc tct gtg ttc cag gcc ttg      168
                 Met Leu Gln Gly His Ser Ser Val Phe Gln Ala Leu
                   1               5                  10

ctg ggg acc ttc ttc acc tgg ggg atg aca gca gct ggg gca gct ctc      216
Leu Gly Thr Phe Phe Thr Trp Gly Met Thr Ala Ala Gly Ala Ala Leu
         15                  20                  25

gtg ttc gta ttc tct agt gga cag agg cgg atc tta gat gga agt ctt      264
Val Phe Val Phe Ser Ser Gly Gln Arg Arg Ile Leu Asp Gly Ser Leu
     30                  35                  40

ggc ttt gct gca ggg gtc atg ttg gca gct tcc tat tgg tct ctt ctg      312
Gly Phe Ala Ala Gly Val Met Leu Ala Ala Ser Tyr Trp Ser Leu Leu
 45                  50                  55                  60

gcc cca gca gtt gag atg gcc acg tcc tct ggg ggc ttc ggt gcc ttt      360
Ala Pro Ala Val Glu Met Ala Thr Ser Ser Gly Gly Phe Gly Ala Phe
                 65                  70                  75

gcc ttc ttc cct gtg gct gtt ggc ttc acc ctt gga gcg gct ttt gtc      408
Ala Phe Phe Pro Val Ala Val Gly Phe Thr Leu Gly Ala Ala Phe Val
             80                  85                  90

tac ttg gct gac ctc ctg atg cct cac ttg ggt gca gca gaa gac ccc      456
Tyr Leu Ala Asp Leu Leu Met Pro His Leu Gly Ala Ala Glu Asp Pro
         95                  100                 105

cag acg gcc ctg gca ctg aac ttc ggc tct acg ttg atg aag aag aag      504
Gln Thr Ala Leu Ala Leu Asn Phe Gly Ser Thr Leu Met Lys Lys Lys
    110                 115                 120

tct gat cct gag ggt ccc gcg ctg ctc ttc cct gag agt gaa ctt tcc      552
```

```
Ser Asp Pro Glu Gly Pro Ala Leu Leu Phe Pro Glu Ser Glu Leu Ser
125                 130                 135                 140

atc cgg ata ggt aga gct ggg ctt ctt tca gac aag agt gag aat ggt      600
Ile Arg Ile Gly Arg Ala Gly Leu Leu Ser Asp Lys Ser Glu Asn Gly
                145                 150                 155

gag gca tat cag aga aag aag gcg gca gcc act ggc ctt cca gag ggt      648
Glu Ala Tyr Gln Arg Lys Lys Ala Ala Ala Thr Gly Leu Pro Glu Gly
            160                 165                 170

cct gct gtc cct gtg cct tct cga ggg aat ctg gca cag ccc ggc ggc      696
Pro Ala Val Pro Val Pro Ser Arg Gly Asn Leu Ala Gln Pro Gly Gly
        175                 180                 185

agc agc tgg agg agg atc gca ctg ctc atc ttg gcc atc act ata cac      744
Ser Ser Trp Arg Arg Ile Ala Leu Leu Ile Leu Ala Ile Thr Ile His
    190                 195                 200

aac gtt cca aag ggt ctc gct gtt gga gtt gga ttt ggg gct ata gaa      792
Asn Val Pro Lys Gly Leu Ala Val Gly Val Gly Phe Gly Ala Ile Glu
205                 210                 215                 220

aag acg gca tct gct acc ttt gag agt gcc agg aat ttg gcc att gga      840
Lys Thr Ala Ser Ala Thr Phe Glu Ser Ala Arg Asn Leu Ala Ile Gly
                225                 230                 235

atc ggg atc cag aat ttc ccc gag ggc ctg gct gtc agc ctt ccc ttg      888
Ile Gly Ile Gln Asn Phe Pro Glu Gly Leu Ala Val Ser Leu Pro Leu
            240                 245                 250

cga ggg gca ggc ttc tcc acc tgg aga gct ttc tgg tat ggg cag ctg      936
Arg Gly Ala Gly Phe Ser Thr Trp Arg Ala Phe Trp Tyr Gly Gln Leu
        255                 260                 265

agc ggc atg gtg gag ccc ctg gcc ggg gtc ttt ggt gcc ttt gcc gtg      984
Ser Gly Met Val Glu Pro Leu Ala Gly Val Phe Gly Ala Phe Ala Val
    270                 275                 280

gtg ctg gct gag ccc atc ctg ccc tac gct ctg gcc ttt gct gcc ggt     1032
Val Leu Ala Glu Pro Ile Leu Pro Tyr Ala Leu Ala Phe Ala Ala Gly
285                 290                 295                 300

gcc atg gtc tac gtg gtc atg gac gac atc atc ccc gaa gcc cag atc     1080
Ala Met Val Tyr Val Val Met Asp Asp Ile Ile Pro Glu Ala Gln Ile
                305                 310                 315

agt ggt aat ggg aaa ctg gca tcc tgg gcc tcc atc ctg gga ttt gta     1128
Ser Gly Asn Gly Lys Leu Ala Ser Trp Ala Ser Ile Leu Gly Phe Val
            320                 325                 330

gtg atg atg tca ctg gac gtt ggc ctg ggc tag ggctgaga cgcttcggac     1179
Val Met Met Ser Leu Asp Val Gly Leu Gly  *
        335                 340

cccgggaaag gccatacgaa gaaacagcag tggttggctt ctatgggaca acaagcttct   1239

ttcttcacat taaaactttt ttccttcctc tcttcttcat ctc                     1282
```

<210> 352
<211> 942

<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (389)..(835)

<400> 352

```
taccggtccg gaattcccgg gtcgacgatt tcgttttatc acttcacagc tctcaaccaa      60

caggggggcct tctttatcct atttaacctc tcagtagagt atttgccttt tgcttattcc    120

tctttccagt caactcttga ttattcatgt aatagaggaa agcattaatg tgaatgatag     180

aagactgccc caccttggcc tctcttctct gccacagcag catccagtca cagaaatgga     240

aacaactagc aggagagctt ctgagtgtag cctttgatgg gcctgatgct gggtaactat     300

gtgcagcctc caacttggtc ccttacctaa gatctttctc ctggcctgcc ctccatgtcg     360

gggagctatc gtcgtcttca agctacag   atg cac atg ttg aat gga gct ctt      412
                                 Met His Met Leu Asn Gly Ala Leu
                                   1               5

ctg gca ttg ctg ttt cct gtg gta aac act cgg ctg ctc ccc ttt gaa      460
Leu Ala Leu Leu Phe Pro Val Val Asn Thr Arg Leu Leu Pro Phe Glu
     10                  15                  20

ttg gag att tac tac att cag cat gtt atg ctc tac gtg gta ccc atc      508
Leu Glu Ile Tyr Tyr Ile Gln His Val Met Leu Tyr Val Val Pro Ile
 25                  30                  35                  40

tac ctg ctt tgg aaa gga ggt gct tac act cca gag ccc ctc agc agt      556
Tyr Leu Leu Trp Lys Gly Gly Ala Tyr Thr Pro Glu Pro Leu Ser Ser
                 45                  50                  55

ttc cgg tgg gct ctt ctc tca act ggc ctc atg ttc ttt tat cac ttc      604
Phe Arg Trp Ala Leu Leu Ser Thr Gly Leu Met Phe Phe Tyr His Phe
             60                  65                  70

agc gtc ttg cag atc ctc ggc ctg gtc acc gaa gtg aat ttg aac aac      652
Ser Val Leu Gln Ile Leu Gly Leu Val Thr Glu Val Asn Leu Asn Asn
         75                  80                  85

atg ctg tgt ccg gcc atc tca gac cca ttc tac ggc ccc tgg tat cgc      700
Met Leu Cys Pro Ala Ile Ser Asp Pro Phe Tyr Gly Pro Trp Tyr Arg
     90                  95                 100

atc tgg gcc tcg gga cac cag act ctc atg acc atg acc cac ggg aag      748
Ile Trp Ala Ser Gly His Gln Thr Leu Met Thr Met Thr His Gly Lys
105                 110                 115                 120

ctg gtc atc ctg ttt tca tac atg gct ggg ccc ttg tgt aaa tat ctg      796
Leu Val Ile Leu Phe Ser Tyr Met Ala Gly Pro Leu Cys Lys Tyr Leu
                125                 130                 135

ctg gat ttg ctc cgg ctt cca gcc aag aaa ata gac tga aggtgcttat       845
Leu Asp Leu Leu Arg Leu Pro Ala Lys Lys Ile Asp  *
            140                 145

ttttttttt tttcctccct gaggaagcaa gtcgtgactt gacttggaga acacccagtt     905

cttgataaaa tcatgggaga gggcaaaaaa aaaaaaa                             942
```

<210> 353
<211> 1123
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (536)..(1066)

<400> 353
aaggatcctt aattaaatta atccccccc cccggggcac cgcgaagcca gcgtgctcag 60

cattcaggga ggcagcggcg ctcgccttca tccaggacaa caaaactgcc atttccgaag 120

agaaaggaaa cggctcgcgg tttctggggt tcccaagcgc gcgcctgcga gggaggccgc 180

gagccgagtc gccacgccca gaaccccgag cccgcccccg cgccacccag cccggccccg 240

ccgccccggc tgcgcacgcg acgccccctc gaggccccgc tcctgcgccc tatttggtca 300

ttcgggggggc aagcggcggg aggggaaacg tgcgcggccg aaggggaagc ggagccggcg 360

ccggctgcgc agaggagccg ctctcgccgc cgccacctcg gctgggagcc cacgaggctg 420

ccgcatcctg ccctcggaac aatgggactc ggcgcgcgag gtgcttgcgg ccgcgctgct 480

cctggggacg ctgcaggtgc tagcgctgct gggggccgcc catgaaagcg cagcc atg 538
Met
1

gcg gca tct gca aac ata gag aat tct ggg ctt cca cac aac tcc agt 586
Ala Ala Ser Ala Asn Ile Glu Asn Ser Gly Leu Pro His Asn Ser Ser
5 10 15

gct aac tca aca gag act ctc caa cat gtg cct tct gac cat aca aat 634
Ala Asn Ser Thr Glu Thr Leu Gln His Val Pro Ser Asp His Thr Asn
20 25 30

gaa act tcc aac agt act gtg aaa cca cca act tca gtt gcc tca gac 682
Glu Thr Ser Asn Ser Thr Val Lys Pro Pro Thr Ser Val Ala Ser Asp
35 40 45

tcc agt aat aca acg gtc acc acc atg aaa cct aca gcg gca tct aat 730
Ser Ser Asn Thr Thr Val Thr Thr Met Lys Pro Thr Ala Ala Ser Asn
50 55 60 65

aca aca aca cca ggg atg gtc tca aca aat atg act tct acc acc tta 778
Thr Thr Thr Pro Gly Met Val Ser Thr Asn Met Thr Ser Thr Thr Leu
70 75 80

aag tct aca ccc aaa aca aca agt gtt tca cag aac aca tct cag ata 826
Lys Ser Thr Pro Lys Thr Thr Ser Val Ser Gln Asn Thr Ser Gln Ile
85 90 95

tca aca tcc aca atg acc gta acc cac aat agt tca gtg aca tct gct 874
Ser Thr Ser Thr Met Thr Val Thr His Asn Ser Ser Val Thr Ser Ala
100 105 110

gct tca tca gta aca atc aca aca act atg cat tct gaa gca aag aaa 922
Ala Ser Ser Val Thr Ile Thr Thr Thr Met His Ser Glu Ala Lys Lys
115 120 125

```
gga tca aaa ttt gat act ggg agc ttt gtt ggt ggt att gta tta acg      970
Gly Ser Lys Phe Asp Thr Gly Ser Phe Val Gly Gly Ile Val Leu Thr
130                 135                 140                 145

ctg gga gtt tta tct att ctt tac att gga tgc aaa atg tat tac tca     1018
Leu Gly Val Leu Ser Ile Leu Tyr Ile Gly Cys Lys Met Tyr Tyr Ser
                150                 155                 160

aga aga ggc att cgg tat cga acc ata gat gaa cat gat gcc atc att     1066
Arg Arg Gly Ile Arg Tyr Arg Thr Ile Asp Glu His Asp Ala Ile Ile
            165                 170                 175

taaggaaatc catggaccaa ggatggaata cagattgatg ctgccctatc aattaat     1123
```

<210> 354
<211> 68
<212> PRT
<213> Homo sapiens

<400> 354

```
Met Ser Pro Leu Leu Pro Leu Ser Tyr Lys Leu Val Leu Cys Phe Pro
  1               5                  10                  15
Thr Pro Asn Gly Val Val Thr His Gly Glu Gln Asn Ala Ser Ser Thr
            20                  25                  30
Asp Ile Glu His Gly Leu Lys Thr Ile Leu Ile Lys Pro Pro Ala Arg
        35                  40                  45
Ile Leu Lys Arg Lys Thr Glu Gly Glu Glu Ser Asn Arg Leu Thr Leu
    50                  55                  60
Pro Thr Thr  *
 65      67
```

<210> 355
<211> 124
<212> PRT
<213> Homo sapiens

<400> 355

```
Met Arg Pro Leu Pro Phe Pro Thr Lys Tyr Leu Gly Lys Phe Ser Trp
  1               5                  10                  15
Lys Trp Phe Phe Leu Ser Leu Asn Ala Gly Asp Tyr His Arg Glu Phe
            20                  25                  30
Ser Thr Val Tyr Phe Glu Phe Glu Ser Phe Gly His Trp Ser Phe Leu
        35                  40                  45
Gln Cys Phe Met Ser Trp Arg Thr Arg Ser Met His Thr His Ile Ser
    50                  55                  60
Val Ser Phe Lys Gly Lys Ile Arg Pro Thr Ser Ala Tyr Leu Leu Leu
 65                  70                  75                  80
Phe Leu Phe Phe Phe Cys Tyr Gly Val Ser Leu Cys Cys Pro Gly Trp
                85                  90                  95
Ser Glu Val Val Ala Arg Ser Arg His Leu Ala Ser Ser Ala Ser Arg
            100                 105                 110
Val His Ala Ile Leu Leu Pro Gln Asn Pro Glu  *
        115                 120         123
```

<210> 356
<211> 343
<212> PRT
<213> Homo sapiens

<400> 356

```
Met Leu Gln Gly His Ser Ser Val Phe Gln Ala Leu Leu Gly Thr Phe
 1               5                  10                  15
Phe Thr Trp Gly Met Thr Ala Ala Gly Ala Ala Leu Val Phe Val Phe
            20                  25                  30
Ser Ser Gly Gln Arg Arg Ile Leu Asp Gly Ser Leu Gly Phe Ala Ala
        35                  40                  45
Gly Val Met Leu Ala Ala Ser Tyr Trp Ser Leu Leu Ala Pro Ala Val
    50                  55                  60
Glu Met Ala Thr Ser Ser Gly Gly Phe Gly Ala Phe Ala Phe Phe Pro
65                  70                  75                  80
Val Ala Val Gly Phe Thr Leu Gly Ala Ala Phe Val Tyr Leu Ala Asp
                85                  90                  95
Leu Leu Met Pro His Leu Gly Ala Ala Glu Asp Pro Gln Thr Ala Leu
            100                 105                 110
Ala Leu Asn Phe Gly Ser Thr Leu Met Lys Lys Lys Ser Asp Pro Glu
        115                 120                 125
Gly Pro Ala Leu Leu Phe Pro Glu Ser Glu Leu Ser Ile Arg Ile Gly
    130                 135                 140
Arg Ala Gly Leu Leu Ser Asp Lys Ser Glu Asn Gly Glu Ala Tyr Gln
145                 150                 155                 160
Arg Lys Lys Ala Ala Ala Thr Gly Leu Pro Glu Gly Pro Ala Val Pro
                165                 170                 175
Val Pro Ser Arg Gly Asn Leu Ala Gln Pro Gly Gly Ser Ser Trp Arg
            180                 185                 190
Arg Ile Ala Leu Leu Ile Leu Ala Ile Thr Ile His Asn Val Pro Lys
        195                 200                 205
Gly Leu Ala Val Gly Val Gly Phe Gly Ala Ile Glu Lys Thr Ala Ser
    210                 215                 220
Ala Thr Phe Glu Ser Ala Arg Asn Leu Ala Ile Gly Ile Gly Ile Gln
225                 230                 235                 240
Asn Phe Pro Glu Gly Leu Ala Val Ser Leu Pro Leu Arg Gly Ala Gly
                245                 250                 255
Phe Ser Thr Trp Arg Ala Phe Trp Tyr Gly Gln Leu Ser Gly Met Val
            260                 265                 270
Glu Pro Leu Ala Gly Val Phe Gly Ala Phe Ala Val Val Leu Ala Glu
        275                 280                 285
Pro Ile Leu Pro Tyr Ala Leu Ala Phe Ala Ala Gly Ala Met Val Tyr
    290                 295                 300
Val Val Met Asp Asp Ile Ile Pro Glu Ala Gln Ile Ser Gly Asn Gly
305                 310                 315                 320
Lys Leu Ala Ser Trp Ala Ser Ile Leu Gly Phe Val Val Met Met Ser
                325                 330                 335
Leu Asp Val Gly Leu Gly  *
            340     342
```

<210> 357
<211> 149
<212> PRT
<213> Homo sapiens

<400> 357

```
Met His Met Leu Asn Gly Ala Leu Leu Ala Leu Leu Phe Pro Val Val
 1               5                  10                  15
Asn Thr Arg Leu Leu Pro Phe Glu Leu Glu Ile Tyr Tyr Ile Gln His
            20                  25                  30
Val Met Leu Tyr Val Val Pro Ile Tyr Leu Leu Trp Lys Gly Gly Ala
        35                  40                  45
Tyr Thr Pro Glu Pro Leu Ser Ser Phe Arg Trp Ala Leu Leu Ser Thr
    50                  55                  60
```

```
Gly Leu Met Phe Phe Tyr His Phe Ser Val Leu Gln Ile Leu Gly Leu
 65                  70                  75                  80
Val Thr Glu Val Asn Leu Asn Asn Met Leu Cys Pro Ala Ile Ser Asp
                85                  90                  95
Pro Phe Tyr Gly Pro Trp Tyr Arg Ile Trp Ala Ser Gly His Gln Thr
            100                 105                 110
Leu Met Thr Met Thr His Gly Lys Leu Val Ile Leu Phe Ser Tyr Met
        115                 120                 125
Ala Gly Pro Leu Cys Lys Tyr Leu Leu Asp Leu Leu Arg Leu Pro Ala
    130                 135                 140
Lys Lys Ile Asp  *
145         148
```

<210> 358
<211> 177
<212> PRT
<213> Homo sapiens

<400> 358

```
Met Ala Ala Ser Ala Asn Ile Glu Asn Ser Gly Leu Pro His Asn Ser
  1               5                  10                  15
Ser Ala Asn Ser Thr Glu Thr Leu Gln His Val Pro Ser Asp His Thr
            20                  25                  30
Asn Glu Thr Ser Asn Ser Thr Val Lys Pro Pro Thr Ser Val Ala Ser
        35                  40                  45
Asp Ser Ser Asn Thr Thr Val Thr Thr Met Lys Pro Thr Ala Ala Ser
    50                  55                  60
Asn Thr Thr Thr Pro Gly Met Val Ser Thr Asn Met Thr Ser Thr Thr
 65                  70                  75                  80
Leu Lys Ser Thr Pro Lys Thr Thr Ser Val Ser Gln Asn Thr Ser Gln
                85                  90                  95
Ile Ser Thr Ser Thr Met Thr Val Thr His Asn Ser Ser Val Thr Ser
            100                 105                 110
Ala Ala Ser Ser Val Thr Ile Thr Thr Thr Met His Ser Glu Ala Lys
        115                 120                 125
Lys Gly Ser Lys Phe Asp Thr Gly Ser Phe Val Gly Gly Ile Val Leu
    130                 135                 140
Thr Leu Gly Val Leu Ser Ile Leu Tyr Ile Gly Cys Lys Met Tyr Tyr
145                 150                 155                 160
Ser Arg Arg Gly Ile Arg Tyr Arg Thr Ile Asp Glu His Asp Ala Ile
                165                 170                 175
Ile
177
```

DECLARATION OF NON-ESTABLISHMENT OF INTERNATIONAL SEARCH REPORT

(PCT Article 17(2)(a), Rule 13*ter*.1(c) and 39)

Applicant's or agent's file reference 21272-019	**IMPORTANT DECLARATION**	Date of mailing (day/month/year) 15 JUN 2001
International application No. PCT/US01/02543	International filing date *(day/month/year)* 25 January 2001 (25.01.2001)	(Earliest) Priority date *(day/month/year)* 25 January 2000 (25.01.2000)
International Patent Classification (IPC) or both national classification and IPC IPC(6): C12P 21/06 and US Cl.: 435/69.1		
Applicant HYSEQ, INC.		

This International Searching Authority hereby declares, according to Article 17(2)(a), that **no international search report will be established** on the international application for the reasons indicated below.

1. ☐ The subject matter of the international application relates to:
   - a. ☐ scientific theories.
   - b. ☐ mathematical theories
   - c. ☐ plant varieties.
   - d. ☐ animal varieties.
   - e. ☐ essential biological processes for the production of plants and animals, other than microbiological processes and the products of such processes.
   - f. ☐ schemes, rules or methods of doing business.
   - g. ☐ schemes, rules or methods of performing purely mental acts.
   - h. ☐ schemes, rules or methods of playing games.
   - i. ☐ methods for treatment of the human body by surgery or therapy.
   - j. ☐ methods for treatment of the animal body by surgery or therapy.
   - k. ☐ diagnostic methods practised on the human or animal body.
   - l. ☐ mere presentations of information.
   - m. ☐ computer programs for which this International Searching Authority is not equipped to search prior art.

2. ☒ The failure of the following parts of the international application to comply with prescribed requirements prevents a meaningful search from being carried out:
   ☐ the description ☒ the claims ☐ the drawings

3. ☒ The failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions prevents a meaningful search from being carried out:
   ☐ the written form has not been furnished or does not comply with the standard.
   ☒ the computer readable form has not been furnished or does not comply with the standard.

4. Further comments:

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Jeffrey S. Lundgren Dorothea Lawrence for Telephone No. (703) 308-0196

Form PCT/ISA/203 (July 1998)